

## MOLECULAR DETECTION SYSTEMS UTILIZING REITERATIVE OLIGONUCLEOTIDE SYNTHESIS

### BACKGROUND OF THE INVENTION

#### Field of the Invention

[0002] The present invention relates generally to the detection and kits for the detection of target molecules and, more particularly, to nucleic acid-based detection assays. Detection is effected by the incorporation into nucleic acids of novel nucleoside and nucleotide analogs containing moieties suitable for detection. These analogs may be incorporated chemically or by biological processes including transcription, abortive transcription and DNA and RNA replication. Thus, such analogs may be useful in any assay utilizing nucleic acids including hybridization, arrays, PCR, abortive transcription, transcription and the like. One nucleic acid based detection assay relies on a method that produces multiple signals from a target molecule by generating multiple copies of detectable oligo- or polynucleotides through reiterative synthesis events on a defined nucleic acid template. The method and kits of the invention may be used in primer extension, terminal transferase additions, reverse transcription, nick translation, ligation, end labelling and to detect mutations, RNA molecules, DNA molecules, pathogens, proteins, methylated CpG sites, diseases or conditions.

## Related Art ,

[0003] The development of various methods for nucleic acid detection and the detection of nucleic acid amplification products has led to advances in the detection, identification, and quantification of nucleic acid sequences in recent years. Nucleic acid detection is potentially useful for both qualitative analyses, such as the detection of the presence of defined nucleic acid sequences, and quantitative analyses, such as the quantification of defined nucleic acid sequences. For example, nucleic acid detection may be used to detect and identify pathogens; detect genetic and epigenetic alterations that are linked to defined phenotypes; diagnose genetic diseases or the genetic susceptibility to a particular disease; assess gene expression during development, disease, and/or in response to defined stimuli, including drugs; as well as generally foster advancements in the art by providing research scientists with additional means to study the molecular and biochemical mechanisms that underpin cellular activity.

[0004] Nucleic acid detection technology generally permits the detection of defined nucleic acid sequences through probe hybridization, that is, the base-pairing of one nucleic acid strand with a second strand of a complementary, or nearly complementary, nucleic acid sequence to form a stable, double-stranded hybrid. Such hybrids may be formed of a ribonucleic acid (RNA) segment and a deoxyribonucleic acid (DNA) segment, two RNA segments, or two DNA segments, provided that the two segments have complementary or nearly complementary nucleotide sequences. Under sufficiently stringent conditions, nucleic acid hybridization may be highly specific, requiring exact complementarity between the hybridized strands. Typically, nucleic acid hybrids comprise a hybridized region of about eight or more base pairs to ensure the binding stability of the base-paired nucleic acid strands. Hybridization technology permits the use of one nucleic acid segment, which is appropriately modified to enable detection, to "probe" for and detect a second, complementary nucleic acid segment with both sensitivity and

specificity. In the basic nucleic acid hybridization assay, a single-stranded target nucleic acid (either DNA or RNA) is hybridized, directly or indirectly, to a labeled nucleic acid probe, and the duplexes containing the label are quantified. Both radioactive and non-radioactive labels have been used.

[0005] However, a recognized disadvantage associated with traditional nucleic acid probe technology is a lower sensitivity when the target sequence is present in low copy number or dilute concentration in a test sample. In many cases, the presence of only a minute quantity of a target nucleic acid must be accurately detected from among myriad other nucleic acids that may be present in the sample. The sensitivity of a detection assay depends upon several factors: the ability of a probe to bind to a target molecule; the magnitude of the signal that is generated by each hybridized probe; and the time period available for detection.

[0006] Several methods have been advanced as suitable means for detecting the presence of low levels of a target nucleic acid in a test sample. One category of such methods is generally referred to as target amplification, which generates multiple copies of the target sequence, and these copies are then subject to further analysis, such as by gel electrophoresis, or by detection of an incorporated nucleotide label, for example. Other methods generate multiple products from a hybridized probe, or probes, by, for example, cleaving the hybridized probe to form multiple products or ligating adjacent probes to form a unique, hybridization-dependent product. Still other methods amplify signals generated by the hybridization event, such as a method based upon the hybridization of branched DNA probes that have a target sequence binding domain and a labeled reporting sequence binding domain.

[0007] There are many variations of target nucleic acid amplification, including, for example, exponential amplification, ligation-based amplification, and transcription-based amplification. An example of an exponential nucleic acid amplification method is the polymerase chain reaction (PCR), which has been disclosed in numerous publications. See, for example, Mullis et al., Cold Spring Harbor Symp. Quant. Biol. 51:263-273

(1986); Mullis et al. U.S. Pat. No. 4,582,788; and Saiki R. et al. U.S. Pat. No. 4,683,194. An example of ligation-based amplification is the ligation amplification reaction (LAR) which is disclosed by Wu et al. in Genomics 4:560 (1989). Various methods for transcription-based amplification are disclosed in U.S. Pat. Nos. 5,766,849 and 5,654,142; and also in Kwok et al., Proc. Natl. Acad. Sci. U.S.A. 86:1173 (1989).

[0008] The most commonly used target amplification method is the polymerase chain reaction (PCR), which consists of repeated cycles of DNA polymerase-generated primer extension reactions. Each reaction cycle includes heat denaturation of the target nucleic acid; hybridization to the target nucleic acid of two oligonucleotide primers, which bracket the target sequence on opposite strands of the target that is to be amplified; and extension of the oligonucleotide primers by a nucleotide polymerase to produce multiple, double-stranded copies of the target sequence. Many variations of PCR have been described, and the method is being used for the amplification of DNA or RNA sequences, sequencing, mutation analysis, and others. Thermocycling-based methods that employ a single primer have also been described. See, for example, U.S. Pat. Nos. 5,508,178; 5,595,891; 5,683,879; 5,130,238; and 5,679,512. The primer can be a DNA/RNA chimeric primer, as disclosed in U.S. Pat. No. 5,744,308. Other methods that are dependent on thermal cycling are the ligase chain reaction (LCR) and the related repair chain reaction (RCR).

[0009] Target nucleic acid amplification may be carried out through multiple cycles of incubation at various temperatures (*i.e.*, thermal cycling) or at a constant temperature (*i.e.*, an isothermal process). The discovery of thermostable nucleic acid modifying enzymes has contributed to rapid advances in nucleic acid amplification technology. See, Saiki et al., Science 239:487 (1988). Thermostable nucleic acid modifying enzymes, such as DNA and RNA polymerases, ligases, nucleases, and the like, are used in methods that are dependent on thermal cycling as well as in isothermal amplification methods.

[0010] Isothermal methods, such as strand displacement amplification (SDA) for example, are disclosed by Fraiser et al. in U.S. Pat. No. 5,648,211; Cleuziat et al. in U.S. Pat. No. 5,824,517; and Walker et al., Proc. Natl. Acad. Sci. U.S.A. 89:392-396 (1992). Other isothermal target amplification methods include transcription-based amplification methods in which an RNA polymerase promoter sequence is incorporated into primer extension products at an early stage of the amplification (WO 89/01050), and a further, complementary, target sequence is amplified through reverse transcription followed by physical separation or digestion of an RNA strand in a DNA/RNA hybrid intermediate product. See, for example, U.S. Pat. Nos. 5,169,766 and 4,786,600. Further examples of transcription-based amplification methods include Transcription Mediated Amplification (TMA), Self-Sustained Sequence Replication (3SR), Nucleic Acid Sequence Based Amplification (NASBA), and variations there of. See, for example, Guatelli et al. Proc. Natl. Acad. Sci. U.S.A. 87:1874-1878 (1990) (3SR); U.S. Pat. No. 5,766,849 (TMA); and U.S. Pat. No. 5,654,142 (NASBA).

[0011] These and other techniques have been developed recently to meet the demands for rapid and accurate detection of pathogens, such as bacteria, viruses, and fungi, for example, as well as the detection of normal and abnormal genes. All of these techniques offer powerful tools for the detection and identification of minute amounts of a target nucleic acid in a sample, although each has strengths and weaknesses.

[0012] One problem, especially for PCR, is that conditions for amplifying the target nucleic acid for subsequent detection and optional quantitation vary with each test, that is, there are no constant conditions favoring test standardization. Further, amplification methods that use a thermocycling process have the added disadvantage of extended lag times which are required for the thermocycling block to reach the "target" temperature for each cycle. Consequently, amplification reactions performed using thermocycling processes require a significant amount of time to reach completion. The various isothermal target amplification methods do not require a thermocycler

and are therefore easier to adapt to common instrumentation platforms. However, the previously described isothermal target amplification methods also have several drawbacks. Amplification according to the SDA methods requires the presence of defined sites for restriction enzymes, which limits its applicability. The transcription-based amplification methods, such as the NASBA and TMA methods, are limited by the need to incorporate a polymerase promoter sequence into the amplification product by a primer.

[0013] Accordingly, there is a need for rapid, sensitive, and standardized nucleic acid signal detection methods that can detect low levels of a target nucleic acid in a test sample. Such a need, as well as others, are met by the inventions of this application.

[0014] A further factor in nucleic acid detection is the choice of label for detecting nucleic acids. Historically, the labels utilized for such purposes have been radioactively labeled with isotopes of hydrogen ( $^3\text{H}$ ), phosphorous ( $^{32}\text{P}$ ), carbon ( $^{14}\text{C}$ ), or iodine ( $^{125}\text{I}$ ). Such radioactive compounds provide useful probes that permit the user to detect, monitor, localize or isolate the labeled molecule of interest. Radioisotopes are useful in many ways; the labelling of molecules with radioisotopes does not usually affect a molecule's physical or chemical properties, labelled molecules can be detected directly (e.g. with film), and the sensitivity to detection is very high.

[0015] There are, however, limitations and drawbacks associated with the use of radioactive compounds. First, since personnel who handle radioactive material can be exposed to potentially hazardous levels of radiation, elaborate safety precautions must be maintained during the preparation, utilization, and disposal of the radioisotopes. Second, radioactive labeled molecules are extremely expensive to purchase and use, in large part due to the cost of equipment and manpower necessary to provide the appropriate safeguards, producer/user health monitoring services and waste disposal programs. Third, radioactive materials may be very unstable and have a limited shelf life, which further increases usage costs. This instability results from radiolytic decomposition due to the destructive effects associated with the decay of the

radioisotope itself and from the fact that many isotopes have half lives of only a few days.

[0016] It has been proposed to chemically label compounds of interest, such as nucleotides and polynucleotides, so as to overcome or avoid the hazards and difficulties associated with such compounds or materials when radioactively labeled. For example, probes consisting of nucleotides labeled with biotin that are capable of being incorporated into double stranded DNA have been prepared. Langer *et al.* (Proc. Natl. Acad. Sci., USA. 78:6633-6637, 1981) describes analogs of dUTP and UTP that contain a biotin molecule bound to the C-5 position of the pyrimidine ring through an alkylamine linker arm. The biotin-labeled nucleotides are efficient substrates for a variety of DNA and RNA polymerases *in vitro*. Polynucleotides containing low levels of biotin substitution (50 molecules or fewer per kilobase) have denaturation, reassociation and hybridization characteristics similar to those of unsubstituted controls. Biotin-labeled polynucleotides, both single and double stranded, are selectively and quantitatively retained on avidin-Sepharose, even after extensive washing with 8M urea, 6M guanidine hydrochloride or 99% formamide. In addition, biotin-labeled nucleotides can be selectively immunoprecipitated in the presence of antibiotin antibody and *Staphylococcus aurea*, Protein A. These unique features of biotin-labeled polynucleotides suggest that they are useful affinity probes for the detection and isolation of specific DNA and RNA sequences. The presence of the biotinylated nucleotide in the hybridized double stranded DNA is detected by utilizing the strong affinity of avidin for biotin, which results in the formation of a stable biotin/avidin complex (binding constant about  $10^{-16}$ ). Attached to the avidin is a biotinylated enzyme, such as biotinylated horseradish peroxidase, which is capable of generating a signal when treated with a suitable reagent. Streptavidin can be used as a substitute for avidin.

[0017] Other methods of labelling nucleotides are known. Nucleotides may be labelled with haptens that bind with a lectin which is, in turn, bound to a detection system to such as an enzyme; with chelating molecules that bind to

metals; or with spin-labelled compounds that may be detected by nuclear magnetic resonance (NMR); or molecules that fluoresce when irradiated with light at specific wavelengths. Nucleic acids may be directly labelled with enzymes such as horseradish peroxidase.

[0018] Modification of nucleotides and nucleotide analogs with labelling molecules can affect the ability of labelled nucleotides to incorporate into nucleic acids, such as by interfering with polymerases; and nucleic acids containing such labelled nucleotides may also exhibit altered properties, such as altered binding to homologous nucleic acids. The nature of the assay, the mode of detection, and the cost of reagents, including the labelled nucleotide also affect the choice of labelled nucleotide. Thus, there remains a need for the use of new types of labelled nucleotides. The present invention provides modified nucleotides for use in methods, assays, molecules and kits.

[0019] All patents, patent publications, and scientific articles cited or identified in this application are hereby incorporated by reference in their entirety to the same extent as if each individual document was specifically and individually indicated to be incorporated by reference in its entirety.

#### SUMMARY OF THE INVENTION

[0020] The invention comprises a nucleoside, nucleotide, oligonucleotide or nucleic acid comprising a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR;

[0021] wherein Nuc is pyrimidinyl or purinyl;

[0022] wherein S is sulfur;

[0023] wherein R is selected from the group consisting of H, haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, a  $C_{6-10}$  aryl,  $C_{6-10}$  ar( $C_{1-6}$ )alkyl,  $C_{6-10}$

- 9 -

arylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> aryloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonylamino(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyl(C<sub>1-6</sub>)alkyl and C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-4 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, C<sub>6-10</sub> arylamino, (C<sub>6-10</sub> aryl)carbonylamino, mono(C<sub>6-10</sub> aryl)aminocarbonyl, di(C<sub>6-10</sub> aryl)aminocarbonyl, mono(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, di(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, N-(C<sub>6-10</sub>)aryl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>6-10</sub> aryl)aminocarbonyl, C<sub>1-6</sub> alkylthio, C<sub>6-10</sub> arylthio, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro, cyano, heteroaryl and saturated or partially unsaturated heterocycle, wherein the heteroaryl and saturated or partially unsaturated heterocycle are independently monocyclic or fused bicyclic and independently have 5 to 10 ring atoms, wherein one or more of the ring atoms are independently selected from the group consisting of oxygen, nitrogen and sulfur; *provided that* (1) when the aryl portion of R is substituted with two substituents, at least one of the substituents is other than nitro, and (2) when R is phenylacetamidomethyl, the phenyl portion is substituted at least once.

[0024] In one embodiment, R is selected from the group consisting of H, haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, phenyl, phenyl(C<sub>1-6</sub>)alkyl,

phenylamino(C<sub>1-6</sub>)alkyl, phenoxy(C<sub>1-6</sub>)alkyl,  
phenyl(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkyloxy (C<sub>1-6</sub>)alkyl,  
phenyl(C<sub>1-6</sub>)alkyl carbonylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)  
alkyl carbonyloxy(C<sub>1-6</sub>)alkyl, benzoylcarbonylamino(C<sub>1-6</sub>)alkyl, benzoyloxy  
(C<sub>1-6</sub>)alkyl, benzoyl(C<sub>1-6</sub>)alkyl and phenyl(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl,  
wherein the aryl portion of each of the preceding groups is optionally  
substituted with 1-2 substituents independently selected from the group  
consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub>  
hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl,  
(C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub>  
alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)  
alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, phenylamino, benzoylamino,  
phenylaminocarbonyl, diphenylaminocarbonyl, phenyl(C<sub>1-6</sub> alkyl)  
aminocarbonyl, di(phenyl(C<sub>1-6</sub> alkyl))aminocarbonyl, N-phenyl-N-(C<sub>1-6</sub> alkyl)  
aminocarbonyl, N-phenyl(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-phenyl  
(C<sub>1-6</sub>)alkyl-N-phenylaminocarbonyl, C<sub>1-6</sub> alkylthio, phenylthio, phenyl(C<sub>1-6</sub>)  
alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro and cyano.

[0025] In one embodiment, R is selected from the group consisting of biotin,  
an enzyme (e.g horseradish peroxidase or alkaline phosphatase), a protein, an  
artificial promoter cassette, streptavidin, a fluorescent moiety, a colorimetric  
moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold,  
silver), a dye, and a nucleic acid cellular uptake group.

[0026] In one embodiment, R is selected from the group consisting of an  
enzyme (e.g horseradish peroxidase or alkaline phosphatase), a protein, an  
artificial promoter cassette, streptavidin, a metal (e.g. gold, silver), a dye, and a  
nucleic acid cellular uptake group.

[0027] In one embodiment, R is selected from the group consisting of  
fluorescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS, and  
Dinitrophenol.

[0028] In one embodiment, R is selected from the group consisting of  
fluorescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS and

- 11 -

dinitrophenol, pyrene, stilbene, coumarine, bimane, naphthalene, pyridoxazole, naphthalamid, NBD, and BODIPY.

[0029] In one embodiment, R is selected from the group consisting of haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, a C<sub>6-10</sub> aryl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> arylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> aryloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonylamino(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyl(C<sub>1-6</sub>)alkyl and C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-4 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, C<sub>6-10</sub> arylamino, (C<sub>6-10</sub> aryl)carbonylamino, mono(C<sub>6-10</sub> aryl)aminocarbonyl, di(C<sub>6-10</sub> aryl)aminocarbonyl, mono(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, di(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, N-(C<sub>6-10</sub>)aryl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>6-10</sub> aryl)aminocarbonyl, C<sub>1-6</sub> alkylthio, C<sub>6-10</sub> arylthio, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro, cyano, heteroaryl and saturated or partially unsaturated heterocycle, wherein the heteroaryl and saturated or partially unsaturated heterocycle are independently monocyclic or fused bicyclic and independently have 5 to 10 ring atoms, wherein one or more of the ring atoms are independently selected from the group consisting of oxygen, nitrogen and sulfur; *provided that* (1) when the aryl portion of R is substituted with two substituents, at least one

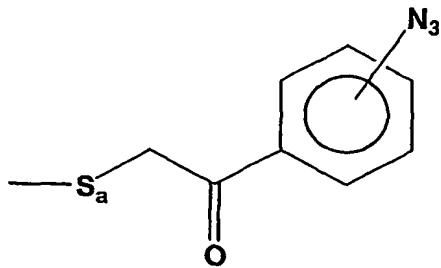
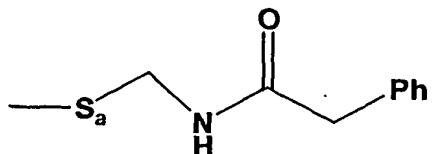
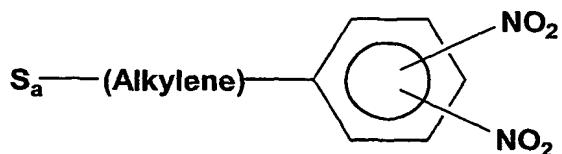
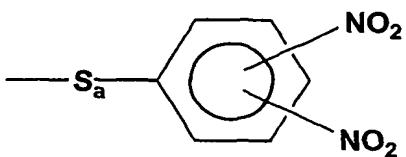
- 12 -

of the substituents is other than nitro, and (2) when R is phenylacetamidomethyl, the phenyl portion is substituted at least once;

[0030] and R is not selected from the group consisting of floorescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS and dinitrophenol, pyrene, stilbene, coumarine, bimane, naphthalene, pyridoxazole, naphthalamide, NBD, and BODIPY.

[0031]

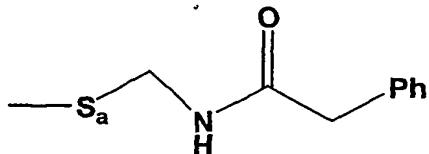
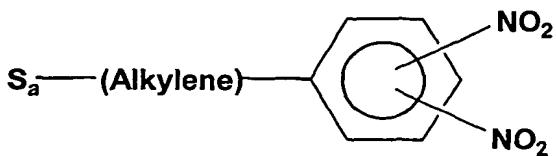
[0032] In another embodiment SR is



- 13 -

wherein a is 1 or 2.

[0033] In another embodiment SR is



wherein a is 1 or 2.

[0034] In some embodiments, SR is substituted with S-L-R, wherein L is a linker having a backbone chain length of 1 to about 8 atoms; wherein the atoms are selected from the group consisting of carbon, sulfur, nitrogen, oxygen and phosphorous.

[0035] In certain embodiments SR is a thiol group that is stable under conditions of nucleic acid synthesis and/or conditions of enzymatic nucleic acid synthesis and is convertible to a reactive thiol after said synthesis.

[0036] In embodiments of the invention, the nucleic acid may be DNA, RNA, or a hybrid species.

[0037] In related embodiments, the invention provides a method for labelling a nucleic acid, wherein the method comprises incorporating into a nucleic acid at least one nucleotide analog that is a 5-S-substituted pyrimidine or 8-S-

substituted purine of the formula NucSR, wherein R is as described herein. The nucleotide analog may be incorporated by chemical coupling; or by an enzyme mediated process selected from the group consisting of: DNA-dependent DNA polymerization; RNA-dependent DNA polymerization; DNA-dependent RNA polymerization; and RNA-dependent RNA polymerization. Suitable enzyme mediated processes for incorporating the nucleotides of the invention include polymerase chain reaction; nick translation; reverse transcription; transcription; primer extension; terminal transferase additions; ligation; end labelling; and abortive transcription.

[0038] In one embodiment, the invention provides for a method of detecting a second nucleic acid of interest, said method involving producing a first molecule of nucleic acid; wherein said first molecule comprises at least one modified nucleotide analog as described herein, mixing said first nucleic acid with a second nucleic acid such that said first and second nucleic acids are associated by Watson Crick base pairing; and detecting the presence of the first nucleic acid. Examples of such methods include Southern blotting; Northern blotting; and Microarray hybridization

[0039] In other embodiments, the invention is a method for detecting multiple reiterated oligonucleotides from a target DNA or RNA polynucleotide, said method comprising:

(a) hybridizing an initiator with a single stranded target polynucleotide

(b) incubating said target polynucleotide and initiator with an RNA-polymerase, elongators and/or a terminator; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein.

(c) synthesizing multiple oligonucleotides from said target polynucleotide, wherein said initiator is extended with one or more elongator nucleotides until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotides; and

detecting or quantifying said reiteratively synthesized oligonucleotide transcripts of a polynucleotide of interest.

[0040] The method of the invention is not required to be performed in any particular order. In one embodiment, the method is carried out in the order (a), (b), (c). In another embodiment, the method is carried out in the order (b), (a), (c). In yet another embodiment, (a) and (b) are performed simultaneously.

[0041] Suitable polymerases for this embodiment include, but are not limited to a DNA-dependent RNA polymerase, an RNA-dependent RNA polymerase and a modified RNA-polymerase, and a primase. In one preferred embodiment, the polymerases comprise an RNA polymerase derived from one of *E. coli*, *E. coli* bacteriophage T7, *E. coli* bacteriophage T3, *S. typhimurium* bacteriophage SP6, and *Thermus aquaticus*.

[0042] In another embodiment, the invention is a method of detecting multiple reiterated oligonucleotides from a target DNA or RNA polynucleotide, said method comprising:

(a) hybridizing an initiator to a single-stranded target polynucleotide;

(b) incubating said target polynucleotide and initiator with a target site probe, an elongator, an RNA-polymerase, and optionally a terminator; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein;

(c) synthesizing an oligonucleotide transcript that is complementary to said target site from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and

detecting or quantifying said reiteratively synthesized oligonucleotide transcripts to about 100 nucleotides and greater than 100 nucleotides.

[0043] The method of the invention is not required to be performed in any particular order. In one embodiment, the method is carried out in the order

(a), (b), (c). In another embodiment, the method is carried out in the order (b), (a), (c). In yet another embodiment, (a) and (b) are performed simultaneously.

[0044] In another embodiment, the invention is a method of detecting multiple reiterated oligonucleotides from a target DNA or RNA polynucleotide, said method comprising:

(a) hybridizing a target site probe to a single-stranded target polynucleotide;

(b) incubating said target polynucleotide and target site probe with an initiator, an elongator, an RNA-polymerase, and optionally a terminator; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein;

(c) synthesizing an oligonucleotide transcript that is complementary to said target site from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and

detecting or quantifying said reiteratively synthesized oligonucleotide transcripts to about 100 nucleotides and greater than 100 nucleotides.

[0045] The method of the invention is not required to be performed in any particular order. In one embodiment, the method is carried out in the order (a), (b), (c). In another embodiment, the method is carried out in the order (b), (a), (c). In yet another embodiment, (a) and (b) are performed simultaneously.

[0046] In one embodiment, the invention is a method for detecting methylated cytosine residues at CpG sites in a target polynucleotide, comprising:

(a) deaminating a single-stranded target DNA sequence under conditions which convert unmethylated cytosine residues to uracil residues while not converting methylated cytosine residues to uracil;

(b) hybridizing an initiator with a said deaminated DNA sequence;

(c) incubating said deaminated target polynucleotide and said initiator with a target site probe, an elongator and/or a terminator, and an

RNA-polymerase, wherein at least one of said initiator, elongator or terminator is modified to enable detection of the CG sites; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein;

(d) synthesizing an oligonucleotide transcript that is complementary to said CG sites from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0047] The method of the invention is not required to be performed in any particular order. In one embodiment, the method is carried out in the order (a), (b), (c), (d). In another embodiment, the method is carried out in the order (b), (a), (c), (d). In yet another embodiment, (a) and (b) are performed simultaneously.

[0048] In one embodiment, the invention is a method for detecting methylated cytosine residues at CpG sites in a target polynucleotide, comprising:

(a) hybridizing an initiator with a DNA sequence;  
(b) incubating said target polynucleotide and said initiator with a target site probe, an elongator and/or a terminator, and an RNA-polymerase, wherein at least one of said initiator, elongator or terminator is modified to enable detection of the CG sites; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein.

(c) synthesizing an oligonucleotide transcript that is complementary to said CG sites from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts;

(d) detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0049] The method of the invention is not required to be performed in any particular order. In one embodiment, the method is carried out in the order (a), (b), (c), (d). In another embodiment, the method is carried out in the order (b), (a), (c), (d). In yet another embodiment, (a) and (b) are performed simultaneously.

[0050] In one embodiment, the invention is a method for detecting the presence or absence of mutations in a target DNA sequence, the method comprising

(a) hybridizing a target site probe to a single-stranded DNA polynucleotide, wherein said DNA polynucleotide may contain a mutation relative to a normal or wild type gene;

(b) incubating said target polynucleotide and target-site probe with an RNA-polymerase, an initiator, elongator and/or terminator; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein;

(c) synthesizing an oligonucleotide transcript from said target polynucleotide that is complementary to a target mutation site, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotides; and

(d) determining the presence or absence of a mutation by detecting or quantifying said reiteratively synthesized oligonucleotides transcribed from said target DNA polynucleotide.

[0051] The method of the invention is not required to be performed in any particular order. In one embodiment, the method is carried out in the order (a), (b), (c), (d). In another embodiment, the method is carried out in the order (b), (a), (c), (d). In yet another embodiment, (a) and (b) are performed simultaneously.

[0052] In one embodiment, the invention is a method for detecting mutations in a target DNA polynucleotide, said method comprising:

(a) immobilizing a capture probe designed to hybridize with said target DNA polynucleotide;

- 19 -

(b) hybridizing said capture probe to said target DNA polynucleotide, wherein said DNA polynucleotide may contain a mutation relative to a normal or wild type gene;

(c) incubating said target polynucleotide with an RNA-polymerase, an initiator, an elongator and/or a terminator; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein;

synthesizing an oligonucleotide transcript that is complementary to a target site from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts;

determining the presence or absence of a mutation by detecting or quantifying said reiteratively synthesized oligonucleotide transcripts from said target DNA polynucleotide.

[0053] The method of the invention is not required to be performed in any particular order.

[0054] In embodiment, the invention is a method for detecting DNA or RNA in a test sample, said method comprising:

(a) hybridizing a single stranded target polynucleotide with an artificial promoter cassette comprising a sequence that hybridizes to the single stranded target polynucleotide, and a region that can be detected by transcription by a polymerase;

(b) incubating said target polynucleotide with an RNA-polymerase, an initiator, an elongator and/or a terminator; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein;

(c) synthesizing an oligonucleotide transcript that is complementary to the initiation start site of the APC, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and

detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0055] The method of the invention is not required to be performed in any particular order. In one embodiment, the method is carried out in the order (a), (b), (c). In another embodiment, the method is carried out in the order (b), (a), (c). In yet another embodiment, (a) and (b) are performed simultaneously.

[0056] In one embodiment, the invention is a method for detecting the presence of pathogens in a test sample, said method comprising:

(a) hybridizing a single stranded target pathogen polynucleotide in said test sample with an artificial promoter cassette comprising a region that can be detected by transcription by a polymerase;

(b) incubating said target polynucleotide and initiator with an RNA-polymerase, an elongator and/or a terminator; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein;

(c) synthesizing an oligonucleotide transcript that is complementary to initiation start site of the APC, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and

(d) determining the presence of a pathogen by detecting or quantifying said reiteratively synthesized oligonucleotide transcripts synthesized from said test sample.

[0057] The method of the invention is not required to be performed in any particular order. In one embodiment, the method is carried out in the order (a), (b), (c), (d). In another embodiment, the method is carried out in the order (b), (a), (c), (d). In yet another embodiment, (a) and (b) are performed simultaneously.

[0058] In one embodiment, the invention is a method for detecting mRNA expression in a test sample, the method comprising:

(a) hybridizing a target mRNA sequence with an artificial promoter cassette comprising a region that can be detected by transcription by a polymerase;

(b) incubating said target mRNA sequence with an RNA-polymerase, an initiator, an elongator and/or terminator; wherein at least one of said initiator, elongator or terminator is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein;

(c) synthesizing an oligonucleotide transcript that is complementary to transcription initiation start site, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotides; and

(d) determining the presence or absence of the mRNA by detecting or quantifying said reiteratively synthesized oligonucleotide transcripts synthesized from said test sample.

[0059] The method of the invention is not required to be performed in any particular order. In one embodiment, the method is carried out in the order (a), (b), (c), (d). In another embodiment, the method is carried out in the order (b), (a), (c), (d). In yet another embodiment, (a) and (b) are performed simultaneously.

[0060] In one embodiment, the invention is a method for detecting an oligonucleotide synthesized from a target DNA sequence, the method comprising:

(a) hybridizing a DNA primer with a single-stranded target DNA sequence;

(b) extending said DNA primer with a DNA polymerase and nucleotides, such that said DNA polymerase reiteratively synthesizes a nucleotide sequence; wherein at least one of said primer and nucleotides is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR, as described herein; and

(c) detecting oligonucleotide comprised of repeat sequences synthesized by said DNA polymerase.

[0061] In methods of the invention, the nucleotide analog is a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR;

[0062] wherein Nuc is pyrimidinyl or purinyl;

[0063] wherein S is sulfur;

[0064] wherein R is selected from the group consisting of H, haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, a C<sub>6-10</sub> aryl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> arylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> aryloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonylamino(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyl(C<sub>1-6</sub>)alkyl and C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-4 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, C<sub>6-10</sub> arylamino, (C<sub>6-10</sub> aryl)carbonylamino, mono(C<sub>6-10</sub> aryl)aminocarbonyl, di(C<sub>6-10</sub> aryl)aminocarbonyl, mono(C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyl)aminocarbonyl, di(C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)aryl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>6-10</sub> aryl)aminocarbonyl, C<sub>1-6</sub> alkylthio, C<sub>6-10</sub> arylthio, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro, cyano, heteroaryl and saturated or partially unsaturated heterocycle, wherein the heteroaryl and saturated or partially unsaturated heterocycle are independently monocyclic or fused bicyclic and independently have 5 to 10 ring atoms, wherein one or more of the ring atoms are independently selected

from the group consisting of oxygen, nitrogen and sulfur; *provided that* (1) when the aryl portion of R is substituted with two substituents, at least one of the substituents is other than nitro, and (2) when R is phenylacetamidomethyl, the phenyl portion is substituted at least once.

[0065] In one embodiment, R is selected from the group consisting of H, haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinkers, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, phenyl, phenyl(C<sub>1-6</sub>)alkyl, phenylamino(C<sub>1-6</sub>)alkyl, phenoxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub> alkyl)carbonylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>alkyl)carbonyloxy(C<sub>1-6</sub>)alkyl, benzoylcarbonylamino(C<sub>1-6</sub>)alkyl, benzoyloxy (C<sub>1-6</sub>)alkyl, benzoyl(C<sub>1-6</sub>)alkyl and phenyl(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-2 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub>alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>) alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, phenylamino, benzoylamino, phenylaminocarbonyl, diphenylaminocarbonyl, phenyl(C<sub>1-6</sub> alkyl) aminocarbonyl, di(phenyl(C<sub>1-6</sub> alkyl))aminocarbonyl, N-phenyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-phenyl(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub>alkyl)aminocarbonyl, N-phenyl(C<sub>1-6</sub>)alkyl-N-phenylaminocarbonyl, C<sub>1-6</sub>alkylthio, phenylthio, phenyl (C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro and cyano.

[0066] In one embodiment, R is selected from the group consisting of biotin, an enzyme (e.g horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, and a nucleic acid cellular uptake group.

[0067] In one embodiment, R is selected from the group consisting of an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, streptavidin, a metal (e.g. gold, silver), a dye, and a nucleic acid cellular uptake group.

[0068] In one embodiment, R is selected from the group consisting of fluorescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS and dinitrophenol, pyrene, stilbene, coumarine, bimane, naphthalene, pyridoxazole, naphthalimid, NBD, and BODIPY.

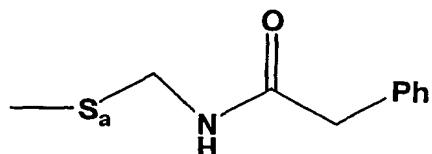
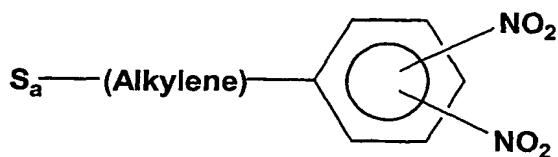
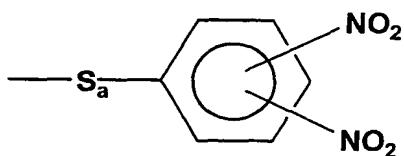
[0069] In one embodiment, R is selected from the group consisting of haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, a C<sub>6-10</sub> aryl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> arylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> aryloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonylamino(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyl(C<sub>1-6</sub>)alkyl and C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-4 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, C<sub>6-10</sub> arylamino, (C<sub>6-10</sub> aryl)carbonylamino, mono(C<sub>6-10</sub> aryl)aminocarbonyl, di(C<sub>6-10</sub> aryl)aminocarbonyl, mono(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, di(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, N-(C<sub>6-10</sub>)aryl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>6-10</sub> aryl)aminocarbonyl, C<sub>1-6</sub> alkylthio, C<sub>6-10</sub> arylthio, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl,

- 25 -

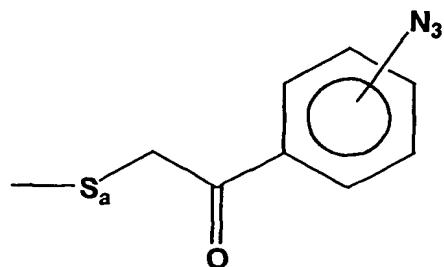
nitro, cyano, heteroaryl and saturated or partially unsaturated heterocycle, wherein the heteroaryl and saturated or partially unsaturated heterocycle are independently monocyclic or fused bicyclic and independently have 5 to 10 ring atoms, wherein one or more of the ring atoms are independently selected from the group consisting of oxygen, nitrogen and sulfur; *provided that* (1) when the aryl portion of R is substituted with two substituents, at least one of the substituents is other than nitro, and (2) when R is phenylacetamidomethyl, the phenyl portion is substituted at least once;

[0070] and R is not selected from the group consisting of fluorescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS and dinitrophenol, pyrene, stilbene, coumarine, bimane, naphthalene, pyridoxazole, naphthalamide, NBD, and BODIPY.

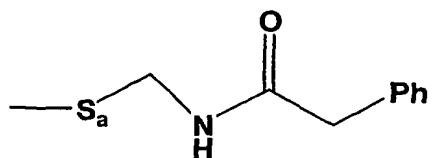
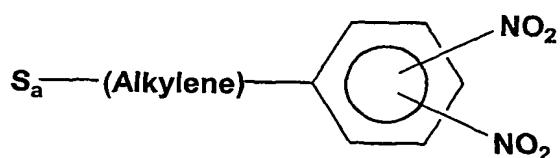
[0071] In another embodiment the SR is



- 26 -



[0072] wherein a is 1 or 2. In another embodiment the SR is



wherein a is 1 or 2.

[0073] In one embodiment R is Coumerin. In one embodiment R is Tetramethylrhodamine. In one embodiment R is Lissamine. In one embodiment R is Alexafluor. In one embodiment R is BODIPY-FL. In one embodiment R is Fluorescene. In one embodiment R is 3'-OMe-SF. In one embodiment R is Texas-Red. In one embodiment R is bis-3'-OMe-ATP. In

one embodiment R is 3'-OMe-SH-ATP. In one embodiment R is TAMARA. In one embodiment R is Biotin. In one embodiment R is H. In one embodiment R is PyMPO.

[0074] In some embodiments, SR is substituted with S-L-R, wherein L is a linker having a backbone chain length of 1 to about 8 atoms; wherein the atoms are selected from the group consisting of carbon, sulfur, nitrogen, oxygen and phosphorous.

[0075] The invention provides methods and compositions for producing multiple detectable signals through reiterative oligonucleotide synthesis reactions on a defined polynucleotide for the detection of target molecules. The invention also provides applications for the reiterative synthesis and detection methods. Important applications of the methods and kits of the invention, include but are not limited to detection of mutations and single nucleotide polymorphisms, DNA molecules, RNA molecules, proteins, pathogens, and detection of pre-cancerous or cancerous mutations and conditions.

[0076] Accordingly, in one aspect, the invention provides a method for synthesizing multiple complementary oligonucleotides from a target DNA or RNA polynucleotide. The method comprises: (a) hybridizing an initiator (nucleoside, mononucleotide, oligonucleotide or polynucleotide) with a single-stranded target polynucleotide (RNA or DNA); (b) incubating said target polynucleotide and initiator with an RNA-polymerase, elongators and/or a terminator nucleotides; (c) synthesizing multiple oligonucleotides from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotides.

[0077] In another aspect, the invention provides a method for detecting multiple reiterated oligonucleotides from a target DNA or RNA polynucleotide. The method comprises: (a) hybridizing an initiator with a single stranded target polynucleotide; (b) incubating said target polynucleotide and initiator with an RNA-polymerase, elongator(s) and/or a terminator

nucleotides; (c) synthesizing multiple oligonucleotides from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released thereby synthesizing multiple reiterative oligonucleotides; and (d) detecting or quantifying said reiteratively synthesized oligonucleotide transcripts of a polynucleotide of interest.

[0078] In a further aspect, the invention provides a method of detecting multiple reiterated oligonucleotides from a target DNA or RNA polynucleotide. The method comprises: (a) hybridizing an target site probe to to a single-stranded target polynucleotide; (b) incubating said target polynucleotide and target site probe with an initiator, an RNA-polymerase, elongators and/or a terminator, wherein said target site probe hybridizes with said target polynucleotide; (c) synthesizing an oligonucleotide transcript that is complementary to said target site from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and (d) detecting or quantifying said reiteratively synthesized oligonucleotide transcripts, wherein said oligonucleotides being synthesized are one of the lengths selected from the group consisting of: about 2 to about 26 nucleotides, about 26 to about 50 nucleotides and about 50 nucleotides to about 100 nucleotides, greater than 100 nucleotides, and greater than 500 nucleotides.

[0079] In a further aspect, the invention provides a method for detecting methylated cytosine residues at CpG sites in a target polynucleotide. The method comprises: (a) hybridizing an initiator with a single stranded target polynucleotide; (b) incubating said target polynucleotide and said initiator with an elongator and/or terminator, an RNA-polymerase; wherein at least one of said initiator, elongator and terminator is modified to enable detection; (c) synthesizing an oligonucleotide transcript that is complementary to said CG site from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby

synthesizing multiple reiterative oligonucleotide transcripts; and (d) detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0080] In a further aspect, the invention provides a method for detecting methylated cytosine residues at CpG sites in a target polynucleotide. The method comprises: (a) deaminating a single-stranded target DNA sequence under conditions which convert unmethylated cytosine residues to uracil residues while not converting methylated cytosine residues to uracil; (b) hybridizing an initiator with a single stranded target polynucleotide; (c) incubating said deaminated target polynucleotide and said initiator with a elongator and/or terminator, an RNA-polymerase, wherein at least one of said initiator, elongator and terminator is modified to enable detection (d) synthesizing an oligonucleotide transcript that is complementary to said CG site from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and (e) detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0081] In still a further aspect, the invention provides a method for detecting methylated cytosine residues at a CpG site in a target gene. The method comprises: (a) deaminating a single-stranded target DNA polynucleotide under conditions which convert unmethylated cytosine residues to uracil residues while not converting methylated cytosine residues to uracil; (b) hybridizing a target site probe with said single stranded target polynucleotide; (c) incubating said target polynucleotide and target site probe with, an initiator, elongator(s) and/or terminator, an RNA-polymerase, wherein said at least one of said initiator, elongator and said terminator are complementary to the CpG site; (d) synthesizing an oligonucleotide transcript that is complementary to said target site from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and (e) detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0082] In still a further aspect, the invention provides a method for detecting the presence or absence of mutations in a target DNA sequence. The method comprises: (a) hybridizing a target site probe to a single-stranded DNA polynucleotide, wherein said DNA polynucleotide may contain a mutation relative to a normal or wild type gene; (b) incubating said target polynucleotide and target-site probe with an RNA-polymerase, a initiator, elongators and/or terminator; (c) synthesizing an oligonucleotide transcript from said target polynucleotide that is complementary to a target mutation site, until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotides; and (d) determining the presence or absence of a mutation by detecting or quantifying said reiteratively synthesized oligonucleotides transcribed from said target DNA polynucleotide.

[0083] In another aspect, the invention provides a method for detecting mutations in a target DNA polynucleotide using a capture probe. The method comprises: (a) immobilizing a capture probe designed to hybridize with said target DNA polynucleotide; (b) hybridizing said capture probe to said target DNA polynucleotide, wherein said DNA polynucleotide may contain a mutation relative to a normal or wild type gene; (c) incubating said target polynucleotide and with an RNA-polymerase, initiator, an elongator and/or terminator; (d) synthesizing an oligonucleotide transcript that is complementary to a target site from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and (e) determining the presence or absence of a mutation by detecting or quantifying said reiteratively synthesized oligonucleotide transcripts from said target DNA polynucleotide.

[0084] In another aspect, the invention provides a method for detecting DNA or RNA in a test sample. The method comprises: (a) attaching a single stranded target polynucleotide with an artificial promoter cassette comprising a sequence that hybridizes to the single stranded target polynucleotide, and a

region that can be detected by transcription by a polymerase; (b) incubating said target polynucleotide with an RNA-polymerase, an initiator, an elongator and/or terminator; (c) synthesizing an oligonucleotide transcript that is complementary to the initiation start site of the APC, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and (d) detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0085] In another aspect, the invention provides a method for detecting the presence of pathogens in a test sample. The method comprises: (a) attaching a target pathogen polynucleotide or protein in said test sample to an artificial promoter cassette comprising a region that can be detected by transcription by a polymerase; (b) incubating said APC-tagged target molecule with an initiator, an RNA-polymerase, an elongator and/or terminator; (c) synthesizing an oligonucleotide transcript that is complementary to the initiation start site of the APC, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and (d) determining the presence of a pathogen by detecting or quantifying said reiteratively synthesized oligonucleotide transcripts synthesized from said test sample.

[0086] In still a further aspect, the invention provides a method for detecting pathogens in a test sample using a capture probe. The method comprises: (a) immobilizing a capture probe designed to interact with a target DNA or RNA polynucleotide in said test sample; (b) mixing said capture probe with a test sample that potentially contains said target polynucleotide; (c) attaching target polynucleotide in said test sample to an artificial promoter cassette comprising a region that interacts with the target pathogen polynucleotide, and a region that can be detected by transcription by a polymerase; (d) incubating said target polynucleotide with an RNA-polymerase, initiator, an elongator and/or terminator; (e) synthesizing an oligonucleotide transcript that is complementary to said initiation transcription start site of APC, wherein said

initiator is extended until transcription terminates and the oligonucleotide is released thereby synthesizing multiple reiterative oligonucleotide transcripts; and (f) determining the presence or absence of a pathogen by detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0087] In still a further aspect, the invention provides a method for detecting mRNA expression in a test sample. The method comprises: (a) attaching a target mRNA sequence to an artificial promoter cassette comprising a region that can be detected by transcription by a polymerase; (b) incubating said target mRNA-APC complex with an RNA-polymerase, an initiator, an elongator and/or terminator; (c) synthesizing an oligonucleotide transcript that is complementary to transcription initiation start site, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotides; and (d) determining the presence or absence of the mRNA by detecting or quantifying said reiteratively synthesized oligonucleotide transcripts synthesized from said test sample.

[0088] In still a further aspect, the invention provides a method for detecting an oligonucleotide synthesized from a target DNA sequence. The method comprises: (a) hybridizing a DNA primer with a single-stranded target DNA sequence; (b) extending said DNA primer with a DNA polymerase and nucleotides, such that said DNA polymerase reiteratively synthesizes a nucleotide sequence; and (c) detecting oligonucleotide comprised of repeat sequences synthesized by said DNA polymerase.

[0089] In still a further aspect, the invention provides a method for producing a microarray. The method comprises: (a) synthesizing multiple oligonucleotide replicates from a target DNA sequence by the methods described herein; and (b) attaching said multiple reiterative oligonucleotide replicates to a solid substrate to produce a microarray of said multiple artificial oligonucleotide replicates.

[0090] In still a further aspect, the invention provides a method for detecting multiple reiterated oligonucleotides from a target DNA or RNA

polynucleotide. The method comprises: (a) incubating a single-stranded target polynucleotide in a mixture comprising an initiator, an RNA-polymerase and optionally additional ribonucleotides; (b) synthesizing multiple oligonucleotide transcripts from said target polynucleotide, wherein said initiator is extended until terminated due to nucleotide deprivation, thereby synthesizing multiple reiterative oligonucleotide transcripts; and (c) detecting or quantifying said reiteratively synthesized oligonucleotides.

[0091] In still a further aspect, the invention provides a method of detecting multiple reiterated oligonucleotides from a target DNA or RNA polynucleotide with a target site probe. The method comprises: (a) incubating a single-stranded target polynucleotide in a mixture comprising an initiator, an RNA-polymerase, a target site probe and optionally additional ribonucleotides, wherein said target site probe and said target polynucleotide hybridize to form a bubble complex comprising a first double-stranded region upstream of a target site, a single-stranded region comprising said target site, and a second double-stranded region near said target site; (b) synthesizing multiple oligonucleotide transcripts from said target polynucleotide, wherein said initiator is extended until terminated due to nucleotide deprivation, thereby synthesizing multiple reiterative oligonucleotides; and (c) detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0092] In still a further aspect, the invention provides a method for detecting methylated cytosine residues at a CG site near or within a target gene. The method comprises:

[0093] (a) deaminating a single-stranded target DNA sequence under conditions which convert unmethylated cytosine residues to uracil residues while not converting methylated cytosine residues to uracil; (b) incubating a single-stranded target polynucleotide in a mixture comprising an initiator, elongators and or terminators, an RNA-polymerase, and a target site probe; (c) synthesizing multiple oligonucleotide transcripts from said target polynucleotide, wherein said initiator is extended until terminated due to nucleotide deprivation, thereby synthesizing multiple reiterative

oligonucleotide transcripts; and (d) detecting or quantifying said reiteratively synthesized oligonucleotides.

[0094] In still a further aspect, the invention provides a method for detecting a target protein in a test sample. The method comprises: (a) covalently attaching the target protein to an artificial promoter cassette (APC) by a reactive APC linker, wherein said APC comprises a region that can be detected by transcription by a polymerase; (b) incubating said target protein with an RNA-polymerase, an initiator, a terminator and optionally additional ribonucleotides; (c) synthesizing an oligonucleotide transcript that is complementary to transcription initiation start site of APC, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and (d) determining the presence or absence of the target protein by detecting or quantifying said reiteratively synthesized oligonucleotide transcripts synthesized from said test sample.

[0095] In still a further aspect, the invention provides a method for detecting cancer. The method comprises: (a) obtaining a sample from a patient in need of detection of a cancer; (b) optionally deaminating the DNA under conditions which convert unmethylated cytosine residues to uracil residues while leaving the methylated cytosine residues unaltered; (c) hybridizing an initiator to a target polynucleotide wherein said initiator is a mononucleoside, mononucleotide, dinucleotide, oligonucleotide, polynucleotide, or an analog thereof; (d) incubating said target polynucleotide and said initiator with an RNA-polymerase and an elongator and/or terminator, wherein at least one of said initiator, terminator, or elongator is modified to enable detection of hybridization to the CG sites; (e) synthesizing an oligonucleotide transcript that is complementary to said CG sites from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released thereby synthesizing multiple reiterative oligonucleotide transcripts; (f) detecting or quantifying said reiteratively

synthesized oligonucleotide transcripts; and (g) comparing the results with those obtained similarly from a control sample.

[0096] In still a further aspect, the invention provides a method for detecting pathogens. The method comprises: (a) obtaining a sample in need of detection of a pathogen; (b) hybridizing a single stranded target pathogen polynucleotide in said sample with an artificial promoter cassette comprising a nucleotide sequence that hybridizes to single stranded target pathogen polynucleotide, and a region that can be detected by transcription by a polymerase; (c) incubating said target polynucleotide and initiator with an RNA-polymerase, an elongator and/or terminator; (d) synthesizing an oligonucleotide transcript that is complementary to initiation start site of the APC, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and (e) determining the presence of a pathogen by detecting or quantifying said reiteratively synthesized oligonucleotide transcripts synthesized from said test sample.

[0097] In another aspect, the method for detecting CG methylation further comprises the incubation of single-stranded target DNA sequence, prior to deamination, with a target-site probe wherein said target site probe and said target DNA sequence form a bubble complex comprising a first double-stranded region upstream of said target CpG site, a single-stranded region comprising said target CpG site, and a second double-stranded region downstream of said target CpG site. In a related aspect, the so-treated DNA is treated with sodium bisulfite to cause deamination of single stranded DNA. In a further related aspect, the conditions for deamination may be milder than those used in the absence of target-site probes. Such conditions include lower temperatures, including below about 50°C, below about 45°C, below about 40°C, and at about 35°C. Conditions also may include a lower period of incubation with deaminating agent, and may include less than about 12 h, less than about 8 h, less than about 6 h, and about 4 h.

[0098] In a related aspect, the invention provides for a method for detecting methylated cytosine residues at CG sites in a target polynucleotide, comprising:

- (a) incubating a single-stranded target DNA sequence with a target-site probe wherein said target site probe and said target DNA sequence form a bubble complex comprising a first double-stranded region near said target CpG site, a single-stranded region comprising said target CpG site, and a second double-stranded region near said target CpG site;
- (b) deaminating target DNA polynucleotide under conditions which convert unmethylated cytosine residues to uracil residues while not converting methylated cytosine residues to uracil;
- (c) hybridizing an initiator to the target polynucleotide;
- (d) incubating said deaminated target polynucleotide and said initiator with an elongator and/or terminator, and an RNA-polymerase, wherein at least one of said initiator, elongator or terminator is modified to enable detection of the CG sites;
- (e) synthesizing an oligonucleotide transcript that is complementary to said CG sites from said target polynucleotide, wherein said initiator is extended until transcription terminates and the oligonucleotide is released, thereby synthesizing multiple reiterative oligonucleotide transcripts; and
- (f) detecting or quantifying said reiteratively synthesized oligonucleotide transcripts.

[0099] The above methods may be employed wherein the target polynucleotide is associated with a gene, is a gene, and/or is a cancer gene.

[00100] In a further aspect, the invention provides a method for determining CG methylation in a sample of interest comprising the use of multiple target specific probes on a single sample to determining the degree of methylation at multiple CG sites. In a related aspect, the determination of the degree of methylation at multiple CG sites is followed by the use of individual target specific probes to determine the degree of methylation at specific CG sites. In

a related aspect, the same sample may be analyzed first with multiple probes; then with single probes.

[00101] The practicing of any of the methods of the invention is not considered to be limited to the particular order listed. For example, in one embodiment, a method may be performed in the order (a), (b), (c), (d). In another embodiment, the method is performed in the order (b), (a), (c), (d). In yet another embodiment, elements of the method may be performed simultaneously.

[00102] The methods of the invention may be practiced with any of the analogs referred herein. As such, methods of the comprise the use of a nucleoside, nucleotide, oligonucleotide or nucleic acid comprising a 5-S-substituted pyrimidine or 8-S-substituted purine of the formula NucSR;

[00103] wherein Nuc is pyrimidinyl or purinyl;

[00104] wherein S is sulfur;

[00105] wherein R is selected from the group consisting of H, haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, a C<sub>6-10</sub> aryl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> arylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> aryloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyl(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyl(C<sub>1-6</sub>)alkyl and C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-4 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub>

alkyl)carbonylamino, C<sub>6-10</sub> arylamino, (C<sub>6-10</sub> aryl)carbonylamino, mono(C<sub>6-10</sub> aryl)aminocarbonyl, di(C<sub>6-10</sub> aryl)aminocarbonyl, mono(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, di(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, N-(C<sub>6-10</sub>)aryl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>6-10</sub> aryl)aminocarbonyl, C<sub>1-6</sub> alkylthio, C<sub>6-10</sub> arylthio, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro, cyano, heteroaryl and saturated or partially unsaturated heterocycle, wherein the heteroaryl and saturated or partially unsaturated heterocycle are independently monocyclic or fused bicyclic and independently have 5 to 10 ring atoms, wherein one or more of the ring atoms are independently selected from the group consisting of oxygen, nitrogen and sulfur; *provided that* (1) when the aryl portion of R is substituted with two substituents, at least one of the substituents is other than nitro, and (2) when R is phenylacetamidomethyl, the phenyl portion is substituted at least once.

[00106] In one embodiment, R is selected from the group consisting of H, haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, phenyl, phenyl(C<sub>1-6</sub>)alkyl, phenylamino(C<sub>1-6</sub>)alkyl, phenoxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub> alkyl)carbonylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub> alkyl)carbonyloxy(C<sub>1-6</sub>)alkyl, benzoylcarbonylamino(C<sub>1-6</sub>)alkyl, benzoyloxy(C<sub>1-6</sub>)alkyl, benzoyl(C<sub>1-6</sub>)alkyl and phenyl(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-2 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)

alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, phenylamino, benzoylamino, phenylaminocarbonyl, diphenylaminocarbonyl, phenyl(C<sub>1-6</sub> alkyl)aminocarbonyl, di(phenyl(C<sub>1-6</sub> alkyl))aminocarbonyl, N-phenyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-phenyl(C<sub>1-6</sub>)alkyl-N-phenylaminocarbonyl, N-phenyl(C<sub>1-6</sub>)alkyl-N-phenylaminocarbonyl, C<sub>1-6</sub> alkylthio, phenylthio, phenyl(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro and cyano.

[00107] In one embodiment, R is selected from the group consisting of biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, and a nucleic acid cellular uptake group.

[00108] In one embodiment, R is selected from the group consisting of an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, streptavidin, a metal (e.g. gold, silver), a dye, and a nucleic acid cellular uptake group.

[00109] In one embodiment, R is selected from the group consisting of fluorescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS, and Dinitrophenol.

[00110] In one embodiment, R is selected from the group consisting of fluorescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS and dinitrophenol, pyrene, stilbene, coumarine, bimane, naphthalene, pyridoxazole, naphthalamid, NBD, and BODIPY.

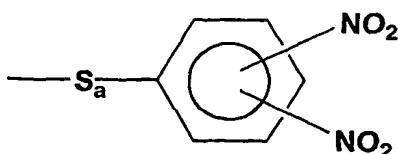
[00111] In one embodiment, R is selected from the group consisting of haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, a C<sub>6-10</sub> aryl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> arylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> aryloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonyloxy(C<sub>1-6</sub>)alkyl,

- 40 -

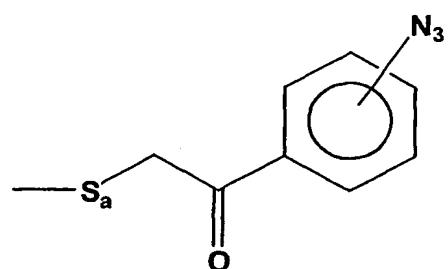
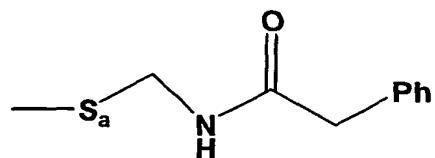
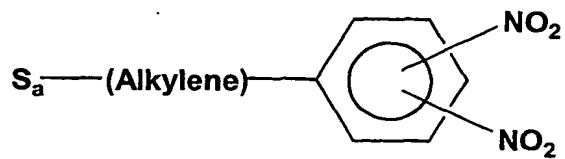
(C<sub>6-10</sub> aryl)carbonylamino(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyl(C<sub>1-6</sub>)alkyl and C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-4 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, C<sub>6-10</sub> arylamino, (C<sub>6-10</sub> aryl)carbonylamino, mono(C<sub>6-10</sub> aryl)aminocarbonyl, di(C<sub>6-10</sub> aryl)aminocarbonyl, mono(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, di(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, N-(C<sub>6-10</sub>)aryl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>6-10</sub> aryl)aminocarbonyl, C<sub>1-6</sub> alkylthio, C<sub>6-10</sub> arylthio, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro, cyano, heteroaryl and saturated or partially unsaturated heterocycle, wherein the heteroaryl and saturated or partially unsaturated heterocycle are independently monocyclic or fused bicyclic and independently have 5 to 10 ring atoms, wherein one or more of the ring atoms are independently selected from the group consisting of oxygen, nitrogen and sulfur; *provided that* (1) when the aryl portion of R is substituted with two substituents, at least one of the substituents is other than nitro, and (2) when R is phenylacetamidomethyl, the phenyl portion is substituted at least once;

[00112] and R is not selected from the group consisting of fluorescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS and dinitrophenol, pyrene, stilbene, coumarine, bimane, naphthalene, pyridoxazole, naphthalamide, NBD, and BODIPY.

[00113] In another embodiment SR is

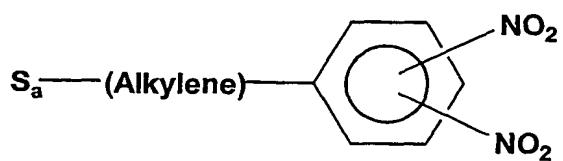


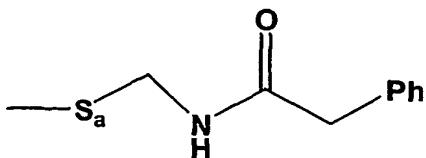
- 41 -



wherein  $a$  is 1 or 2.

[00114] In another embodiment SR is





wherein a is 1 or 2.

[00115] In one embodiment R is Coumerin. In one embodiment R is Tetramethylrhodamine. In one embodiment R is Lissamine. In one embodiment R is Alexafluor. In one embodiment R is BODIPY-FL. In one embodiment R is Fluorescene. In one embodiment R is 3'-OMe-SF. In one embodiment R is Texas-Red. In one embodiment R is bis-3'-OMe-ATP. In one embodiment R is 3'-OMe-SH-ATP. In one embodiment R is TAMARA. In one embodiment R is Biotin. In one embodiment R is H. In one embodiment R is PyMPO.

[00116] In some embodiments, SR is substituted with S-L-R, wherein L is a linker having a backbone chain length of 1 to about 8 atoms; wherein the atoms are selected from the group consisting of carbon, sulfur, nitrogen, oxygen and phosphorous.

[00117] In certain embodiments SR is a thiol group that is stable under conditions of nucleic acid synthesis and/or conditions of enzymatic nucleic acid synthesis and is convertible to a reactive thiol after said synthesis.

[00118] The present invention also provides kits for conducting the oligonucleotide synthesis and detection methods described herein. In one aspect, for example, the invention provides reagent containers, which contain various combinations of the components described herein. These kits, in suitable packaging and generally (but not necessarily) containing suitable instructions, contain one or more components used in the oligonucleotide synthesis and detection methods. The kit may also contain one or more of the following items: enzymes, initiators, primers, buffers, nucleotides, peptides,

proteins, control DNA, antibodies, streptavidin, and biotin. The kit may also contain reagents mixed in appropriate amounts for performing the methods of the invention. The reagent containers preferably contain reagents in unit quantities that obviate measuring steps when performing the subject methods.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[00119] FIG. 1: Artificial Promoter Cassettes. Artificial Promoter Cassettes (APC) are regions of nucleic acid that form a polymerase binding site and can be attached to other macromolecules through interaction with a specific nucleic acid sequence or other molecule, which is termed the APC linker. APC linkers can be attached to target nucleic acids (DNA or RNA) by hybridization to complementary sequences on either the template or non-template strands of the target nucleic acid. An APC Linker can also hybridize to a complementary sequence placed on any target molecule, such as a protein, for detection of molecules that bind to said protein. APC-linkers may contain maleimide or other SH-reactive groups for attachment to SH groups on nucleotide analogs. Multiple detectable oligonucleotides are generated by polymerase bound to the Artificial Promoter Cassette. In this figure, the APC depicted contains two regions of essential complementarity (A, A' and C, C'), which are separated by a "bubble region." In this schematic, the "bubble region" is generated because regions of the two strands are non-complementary (B, and E). Alternatively, the APC may have two completely complementary strands. Upon binding of the RNA polymerase, the DNA strands separate, which leads to the formation of the "bubble region."

[00120] Regions A, B, and C are on one strand. Regions C', E, and A' are on the complementary strand. The APC may be made from two separate strands (ABC and C'EA') or all 6 regions may be on a single polynucleotide, in which regions C and C' are separated by a linker region D, which can modified to be as long as needed. Linker region D may serve only to join C and C' or the sequence of region D may serve as a binding site for other factors that may

enhance abortive transcription, such as transcription roadblock proteins, including but not limited to EcoRI QIII mutant, the lac repressor and other RNA polymerases. The linker region D may be designed for a single road block protein, or multiple roadblock proteins. The length of linker region D will depend on the function of the linker region.

[00121] FIG. 2: Signal Generation by Reiterative oligonucleotide synthesis. A signal is generated by the enzymatic incorporation of one or more nucleotide analogs into multiple (n) highly similar or identical oligonucleotide products. Under appropriate conditions, RNA oligonucleotides can be made from nucleic acid templates in the absence of a promoter. An initiator may be comprised of one or more nucleosides, nucleoside analogs, nucleotides, or nucleotide analogs. The initiator may contain one or more covalently joined nucleotides, including but not limited to, 1-25 nucleotides, 26-50 nucleotides, 51-75 nucleotides, 76-100 nucleotides, 101-125 nucleotides, and 126-150 nucleotides, 151-175 nucleotides, 176-200 nucleotides, 201-225 nucleotides, 226-250 nucleotides and more than 250 nucleotides, and may contain one or more functional R groups. The initiator (n copies) can be elongated directly with n copies of a terminator to end chain elongation or n copies of other elongator nucleotides (Y positions) may be incorporated between the initiator and the terminator to form a longer oligonucleotide. The terminator may contain a second functional group.  $N_I$  = Initiating mononucleotide or oligonucleotide analog,  $N_E$  = Elongating mononucleotides or analog,  $N_T$  = Terminating mononucleotide or analog,  $Z = x + y$ ;  $R_1 = H, OH$ , or reporter group;  $R_2 = H, OH$ , or reporter group;  $N$  = deoxy or ribonucleotides; Polymerase = RNA-dependent or DNA-dependent RNA polymerase. DNA or RNA may be attached to other molecules, such as proteins

[00122] FIG. 3: Some nucleotides that can be elongators or terminators, for use in the methods and compositions of the invention. Nucleotide analogs that may be included at internal or 3' terminal positions in oligonucleotides are shown. All of these analogs can be converted to terminators simply by

- 45 -

replacement of the 3' OH group. These analogs may also function as terminators with the 3' OH group under particular reaction conditions.

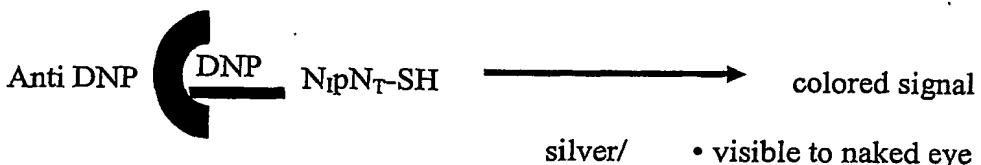
[00123] FIG. 4: Other fluorescent groups that may be R. Such groups would be useful in the methods and compositions of the invention after addition onto the 5-S- position in a pyrimidine or the 8-S- position in a purine.

[00124] FIG. 5: Dinucleotide synthesis via abortive initiation on single-stranded DNA or RNA. Single stranded (ss) nucleic acid is DNA or RNA. Polymerase is a DNA-dependent or RNA-dependent RNA polymerase.  $N_I$  = 3'-OH nucleoside or nucleotide initiator;  $N_T$  = 5'-triphosphate nucleotide or nucleotide analog terminator.  $R_1$  may be on the 5' phosphate group, the 2' position of the sugar, or on the purine or pyrimidine base.  $R_2$  may be on the pyrimidine or purine base or 2' or 3' position of the sugar/ribose or deoxyribose.  $R_1$  = H, OH, and/or any reporter group or reporter group precursor, as described herein.  $R_2$  = H, OH, and/or any reporter group or reporter group precursor, as described herein. Signal may be any signal that can be detected, and includes but is not limited to fluorescence, fluorescence resonance energy transfer (FRET), or colorimetric. As one example,  $R_1$  may be AEDANS, and  $R_2$  may be Fluorescein. Signal is generated by FRET from  $R_1$  to  $R_2$ .

[00125] FIG. 6: Signal generation via dinucleotide production. Oligonucleotides can be synthesized that contain one R group on the initiator nucleotide and another on the terminator nucleotide, such that the R groups have different functions. For example, if  $R_1$  is biotin, it can be used for oligonucleotide product immobilization and  $R_2$  allows for signal production.

[00126] Example 1:  $R_1$  tag = biotin,  $R_2$  tag = fluorescein: detection of fluorescein emission

[00127] Example 2:  $R_1$  tag = DNP,  $R_2$  tag = reactive thiol



- 46 -

gold development	without irradiation
	• point of care devices

[00128] FIG 7: Signal generation for FRET detection by abortive initiation.

Oligonucleotides can be reiteratively synthesized that contain 2 or more nucleotides and have different R groups, one at or near each end of the oligonucleotide product made during transcription. Energy transfer between the two R groups on the substrates can only occur after they are brought into proximity during template-directed oligonucleotide synthesis by enzymatic phosphodiester bond formation between the labeled initiator and the labeled terminator nucleotides. The R<sub>1</sub> donor group on N<sub>I</sub> can be excited by irradiating the sample with light of wavelength of  $\lambda_{1A}$ , where  $\lambda_{1A}$  is the absorption maximum of group R<sub>1</sub>. The excited R<sub>1</sub> donor group emits light of wavelength  $\lambda_{1E}$ , where  $\lambda_{1E}$  is the emission maximum for group R<sub>1</sub> and also a wavelength for absorption by group R<sub>2</sub> ( $\lambda_{2A}$ ). The acceptor R<sub>2</sub> group on N<sub>T</sub> absorbs light of wavelength  $\lambda_{1E}/\lambda_{2A}$  that was emitted by the excited R<sub>1</sub> donor group on N<sub>I</sub>. The excited acceptor R<sub>2</sub> group on N<sub>T</sub> emits light of wavelength  $\lambda_{2E}$ , which is detected and quantified. Similarly, R<sub>2</sub> may be an energy donor to R<sub>1</sub>, with emission from R<sub>1</sub> detected. In the absence of target-associated template, no oligonucleotide is synthesized.

[00129] FIG. 8: Trinucleotide energy transfer. Labeled oligonucleotide synthesis is initiated with a labeled dinucleotide initiator. The label may be on either the 5' nucleotide (R<sub>1</sub>) or the 3' nucleotide (R<sub>2</sub>) of the dinucleotide initiator. The initiator is elongated with a labeled (R<sub>3</sub>) 5'-nucleosidetriphosphate terminator nucleotide analog. Detection via energy transfer can be adjusted to utilize R<sub>1</sub> or R<sub>2</sub> with R<sub>3</sub>, as shown. In the absence of nucleic acid template-directed phosphodiester bond formation between the initiator and terminator, the R groups remain sufficiently separated that no energy transfer is detected. In this example, the amount of energy emitted as  $\lambda_{3E}$  is directly proportional to the amount of template-associated target present.

Similarly, the R groups may be varied for other applications, as demonstrated in Figure 6.

[00130] FIG. 9: Target Site Probe. An RNA polymerase can be directed to specific nucleotide positions (sites) in target nucleic acids by the hybridization of a gene-specific or region-specific Target Site Probe (TSP). The target site is a nucleotide position in the DNA to be analyzed for sequence (as in detection of single nucleotide polymorphisms) or structure (as in assessing the methylation status of a specific nucleotide), and it is located on the template strand of the target sequence near the junction of regions E and C' in the target sequence. The TSP contains a region of homology to the target nucleic acid (Region A) which begins approximately 8-14 nucleotides and ends approximately 15-35 nucleotides upstream of the target site nucleotide. A second region of the TSP is designed to be non-complementary to the 8-14 nucleotides immediately upstream of the target site (Region B), so that a melted "bubble" region forms when the TSP hybridizes to the target nucleic acid. The TSP contains a third region (Region C) which is essentially complementary to the 5-25 nucleotides immediately downstream of the target site nucleotide. RNA polymerase will bind to the bubble complex such that transcription will start near the E/C' junction and will move downstream into the C/C' hybrid.

[00131] FIG. 10: Methylation of CpG Islands in DNA. The human genome has a 4-5 fold lower frequency of CpG dinucleotides than expected given the overall frequency of C and G in human DNA. A large fraction of CpG sequence is distributed into clusters known as CpG islands. These sequence patterns are between 300-3000 nucleotides long and overlap with about 60% of all human promoters. The remaining CpG dinucleotides outside of CpG islands generally contain methylated C. CpG methylation outside of CpG islands stabilize the genome by inactivating the expression of parasitic DNA, and independently play an essential role in development. Changes in the methylation status of cytosine in CpG islands are early events in many cancers and permanent changes found in many tumors. These CpG islands are found

in the regions next to genes that determine whether the gene is "ON" or "OFF". Many genes that are important for preventing cancer, such as tumor suppressor genes, need to be "ON" for cells to grow normally. Cellular enzymes can add methyl groups (methylation) to the C residues in these CpG islands. This methylation results in the shutting "OFF" of these genes. When tumor suppressor genes are shut "OFF", the cell no longer makes the proteins that they encode, and the cell begins to grow without control checkpoints. This is one of the early events that can lead to cell "transformation" and the progression of cancer.

[00132] FIG. 11: Deamination conversion of unmethylated cytosine groups in DNA. Deamination converts unmethylated C to U. Methylated C groups, such as those in CpG islands that regulate eukaryotic genes, are resistant to deamination and remain as C in the product DNA. If 100% deamination occurs, methylated DNA will still contain CpG doublets, whereas unmethylated DNA will contain no cytosine and will now contain UpG where CpG doublets were before deamination. This difference in DNA sequence can be used to distinguish between methylated and unmethylated DNA by transcription because the two DNAs encode different dinucleotides.

[00133] FIG. 12: Detection of methylation using dinucleotide synthesis. Dinucleotide synthesis can be used to assess the overall methylation state of DNA. In the presence of RNA polymerase, CTP or a CTP analog ( $R_1$ -C-OH), and GTP or a GTP analog ( $R_1$ -CpG- $R_2$ ), the deaminated methylated DNA template will produce n copies of a labeled dinucleotide product, where n is proportional to the number of methylated CpG dinucleotides in the starting DNA. The deaminated unmethylated DNA template can produce no dinucleotide with these substrates because the template no longer encodes "C" at any position.

[00134] If  $R_1$  and  $R_2$  are appropriately labeled, the dinucleotide will produce a signal that is proportional to the number of methylated CpG sites. For example, if  $R_1$  is a fluorescent energy donor or acceptor that is compatible with a second donor or acceptor,  $R_2$ , a signal will be detected by fluorescent

resonance energy transfer (FRET) between  $R_1$  and  $R_2$  only when the two groups are brought into proximity after incorporation into the dinucleotide in an enzymatic, template-dependent reaction. The reiterative synthesis of these dinucleotides during abortive transcription results in multiple signals from each CpG target and can be used to assess the methylation level of the DNA.

[00135] Similarly, abortive synthesis of trinucleotides by transcription initiation with labeled dinucleotides that end in C (ApC, CpC, GpC, UpC) and termination with labeled GTP can be used to produce signal from the deaminated methylated template, but not the deaminated unmethylated template. This trinucleotide synthesis approach may be expanded by the addition of a site-specific oligonucleotide to allow assessment of the methylation status of a specific CpG site, rather than the entire island, as illustrated in Figure 13.

[00136] FIG. 13: Assessing methylation status of specific CpG sites in CpG islands by abortive initiation. Target site probes can be used to examine the methylation status of specific CpG islands in specific genes. In the deaminated methylated DNA, the dinucleotide CpG is encoded by the template at the 3 methylated sites 1, 3 and 4, but not by the unmethylated site 2. To specifically determine if Site 3 is methylated and if so, to what extent, position (C21) can be targeted with a Target Site Probe, as described in Figure 9. The template C in question is positioned at the junction of the bubble region and the downstream duplex so that it encodes the next incorporated nucleotide for appropriately primed RNA polymerase that binds to the bubble region. If a labeled initiator  $R_1-N_xpC-OH$  is used, where  $N_x$  may be C for a dinucleotide CpC initiator or  $N_x$  may be CpC for a trinucleotide initiator, etc., the initiator can be elongated with a labeled GTP analog  $pppG-R_{2G}$  to form a trinucleotide  $R_1N_xCpG-R_2$ . Similarly, if the C in question was not methylated, the position will now be a U and will encode nucleotide A. If an ATP analog  $pppA-R_{2A}$  is present, it will be incorporated opposite positions where the C was not methylated. If the GTP analog is labeled with group  $R_{2G}$ , which is an energy acceptor from the R group on the initiator,  $R_1$ , then the amount of

- 50 -

$R_1N_xCpGR_2G$ , which will be proportional to the amount of methylated C present at that position, can be quantified by measuring the emission from  $R_{2G}$  at wavelength  $\lambda_{2GE}$ . The similar situation exists for incorporation of the ATP analog and measurement of the emission from its R group, also an energy acceptor from the initiator  $R_1$ . By determining the ratio of the magnitude of emission from the GTP analog to the total emission from both the ATP and GTP analogs, the site can be assigned a methylation index M. If all of the Cs at that position are methylated, M = 1. If none of the site is methylated, M = 0.

[00137] FIG. 14: Genes with altered methylation in cancer. Forty-nine genes with methylation changes associated with cancer initiation and progression are plotted versus 13 cancers. An oval indicates an abnormal methylation status for a gene, coded by cancer type. Cancer is actively prevented through the expression of close to 100 tumor-suppressor genes that regulate the cell-division cycle. CpG methylation potentially is a powerful biomarker for cancer detection. Examination of the promoters of tumor suppressor genes from tumor biopsies suggests that CpG methylation is common enough to equal the impact of mutagenesis in tumor promotion. At least 60 tumor suppressor and repair genes are associated with abnormally high levels of CpG methylation across virtually all of the common tumor types. In virtually all cases, defective expression of tumor suppressor genes begins at an early stage in tumor progression. Detection of these early methylation events before advanced symptoms appear should improve the chances that a cancer will be treated while it is highly curable. CpG methylation patterns are frequently biased to particular genes in particular types of cancers. Therefore, it should be possible to develop methylation signatures for common cancers, indicating both cancer type and stage. Data on the methylation status of multiple promoters could give clues as to the location of a tumor in cases where several organs can contribute to a sample. For example, shed bladder, kidney or prostate cells can populate a urine sample. Tumors from each of these tissues are frequently associated with distinct combinations of CpG island methylation.

[00138] FIG. 15: Single nucleotide polymorphism detection by abortive oligonucleotide synthesis. The identity of a nucleotide at a specific position can also be determined by abortive initiation in the presence of target nucleic acid and a position-specific Target Site Probe. This can be applied to SNP identification by initiating transcription with an oligonucleotide complementary to the DNA upstream from the SNP site. For example for synthesis of a trinucleotide, the dinucleotide initiator would be complementary to the known nucleotides at positions -1 and -2, relative to the SNP site.

[00139] FIG. 16: Detection and identification of single nucleotide polymorphisms (SNPs) by abortive transcription. The identity of a specific DNA nucleotide (A,C,G,T/dU) can be identified by abortive transcription with the use of a Target Site Probe (TSP). For example, to determine whether a DNA contains a normal nucleotide (wild type) or a mutant nucleotide (point-mutation, single nucleotide polymorphism/ SNP), a gene-specific TSP can be added to target DNA (or amplification/replication product) such that the SNP position corresponds to the last nucleotide in the C/C' hybrid at the junction of the downstream duplex and the bubble region. A labeled initiator oligonucleotide ( $R_1NI-OH$ ) that is complementary to the region upstream of the SNP site can be elongated by an RNA polymerase to add the next encoded nucleotide, corresponding to the SNP. The labeled terminators (pppN<sub>T</sub>-R<sub>2</sub> or pppU-R<sub>2U</sub>, pppA-R<sub>2A</sub>, pppC-R<sub>2C</sub>, pppG-R<sub>2G</sub>) can each be labeled with different R groups, for example, R<sub>2A</sub>, R<sub>2C</sub>, R<sub>2G</sub> and R<sub>2U</sub> could each be resonance energy acceptors from R<sub>1</sub>, with each emitting light with a different detectable wavelength.

[00140] FIG. 17: Signal Generation from artificial promoter. An Artificial Promoter Cassette (APC) consists of one or more oligonucleotides or polynucleotides that together create a specific binding site for an RNA polymerase coupled to a linker region (APC linker) for attachment to target molecules (DNA, RNA, Protein). The APC may contain an artificial promoter, or it may contain the promoter for a specific RNA polymerase. For example, trinucleotide or tetranucleotide products that could be generated from with a

common phage RNA polymerase can be made with a labeled GpA or GpApA initiator and a labeled pppG or pppA terminator.

[00141] FIG. 18: Detection of nucleic acids by abortive transcription. For detection of nucleic acids, such as DNA or RNA associated with specific diseases or with viral and bacterial pathogens, one can either detect the nucleic acid directly or after replication or primer extension. In the first case, the APC linker in the Artificial Promoter Cassette would be designed to be complementary to a known DNA or RNA sequence of the target nucleic acid. Alternatively, one or more copies of the target DNA or cDNA copies of target RNA can be made by primer extension or reverse transcription initiated with primers containing a universal APC linker sequence at the 5' end. In either case, the target DNA or RNA can be retrieved from the sample by attachment to a solid support, for example, to which an oligonucleotide that contains a second target-specific sequence, which is termed a "capture sequence," has been attached via any number of immobilization tags, including but not limited to biotin, hexahistidine or any other hapten. Once attached, abortive transcription is initiated by addition of a polymerase and the appropriate labeled nucleotides, which results in signal generation, as previously described.

[00142] FIG. 19: Detection of mRNA by Abortive Transcription. An Artificial Promoter Cassette for detection of mRNA will contain as its APC linker an oligo T tail. This tail is complementary to the poly A tail found at the 3' end of eukaryotic mRNAs and will be used for attachment of the APC to the target mRNA. The target mRNA can be retrieved from a sample by attachment to an immobilized capture probe containing a capture sequence, which is complementary to some region of the target mRNA.

[00143] FIG. 20: Detection of proteins or other haptens/antigens with abortive transcription. Signal generation via abortive initiation from an Artificial Promoter Cassette can be used to detect other molecules, such as proteins. For example, an APC linker sequence can be prepared to which thiol-reactive or amine-reactive protein crosslinking agents R will be covalently attached. The

reactive APC linker will be added to the target protein, which may be purified or in a complex mixture (such as a cell lysate), and the APC linker will be covalently attached to the target protein via modification of protein thiol and/or amine groups. The labeled protein can then be immobilized utilizing a target-specific probe (such as an antibody). The Artificial Promoter Cassette is then attached via the APC linker, and signal is generated, as previously described.

[00144] FIG. 21: Enhanced detection of molecular targets via abortive transcription on APC particles. Even greater detection sensitivity can be achieved with the use of particles to which multiple copies, including tens, hundreds, thousands, tens of thousands or even more of the Artificial Promoter Cassette (APC) have been attached. The sphere will also contain a linker that will be specific for binding to a group that can be attached to the target molecule. For example, streptavidin (SA) can be attached to the APC particles and biotin to the target molecule, which can then be immobilized via interaction with a target-specific capture probe. Once the APC particles interact with the target, for example via the SA-biotin interaction, polymerase and labeled nucleotides can be added for signal generation, as described.

[00145] FIG. 22: Coating of DNA or RNA targets with APC particles for ultra-sensitive detection or molecular imaging. An alternative method for the ultra-sensitive detection or visualization of target DNA or RNA can be achieved by reverse transcription of target RNA or copying (single copy or amplification) of target DNA in the presence of probe labeled dNTP analogs. As an example, 5-SH-dUTP can be incorporated at very high frequency in DNA molecules, which can then be immobilized and further modified with other groups, such as biotin. To this, APC particles can be added, as described in Figure 21, each of which will interact with a nucleotide analog on the target. This essentially coats the target DNA or RNA with APC particles capable of generating multiple oligonucleotide products for a variety of methods of molecular detection.

[00146] FIG 23. Detection of telomerase activity with reiterative oligonucleotide synthesis. Reiterative oligonucleotide synthesis with DNA polymerases can also be used for signal generation, however, the product oligonucleotides need not be released, but may be joined tandemly in the product. As an example, telomerase activity can be detected by immobilizing a telomerase-specific probe to a solid matrix to capture cellular telomerase, which carries its own RNA template for DNA synthesis. For example, with human telomerase, the RNA template on the enzyme encodes the DNA sequence GGGTTA. The capture probe may contain the sequence GGGTTA, which will be added reiteratively to the end of the telomerase capture probe, if telomerase is present in the sample. Signal generation can be achieved in several ways, one of which involves including one or more reporter tagged dNTPs in the synthesis reaction to produce a product that has multiple R<sub>1</sub> groups attached along the backbone of the DNA product. For detection, this product can then be hybridized to a complementary probe containing nucleotides with a second R group (R<sub>2</sub>) attached that will hybridize to the R<sub>1</sub> labeled product. This will bring the R<sub>1</sub> and R<sub>2</sub> groups together for signal generation via FRET from between R<sub>1</sub> and R<sub>2</sub>, or via other methods. Alternatively, telomerase may incorporate 2 labeled nucleotides in the product DNA and look for energy transfer between the 2 labeled nucleotides in the single strand of DNA.

[00147] FIG. 24. Synthesis of a dye labeled initiator. 5'EADANS-S-CMP was synthesized from the conjugation of IAEDANS and  $\alpha$ -S-CMP. The scanned image of the thin layer chromatography plate shows the control IAEDANS and the IAEDANS labeled product. Lane 1: Cytidine-5'-O-(1-Thiomonophosphate); Lane 2: Cytidine-5'-O-(1-Thiotriphosphate); Lane 3: 5-(((2-iodoacetyl)amino)ethyl)amino)naphthalene-1-sulfonic acid (1,5-IAEDANS); Lane 4: Cytidine-5'-O-(1-Thiotriphosphate) and (1,5-IAEDANS); Lane 5: Cytidine-5'-O-(1-Thiomonophosphate) and (1,5-IAEDANS); Lane 6: Adenosine-5'-O-(1-Thiomonophosphate); Lane 7: Adenosine-5'-O-(1-Thiomonophosphate) and (1,5-IAEDANS); Lane 8:

(1,5-IAEDANS); Lane 9: Cytidine-5'-O-(1-Thiotriphosphate); Lane 10: Cytidine-5'-O-(1-Thiomonophosphate). Lanes 4, 5, and 7 also contain 1 U of *E. coli* RNA polymerase, Buffer T and 150 ng of denatured pBR322

[0100] FIG. 25. Abortive Transcription Initiation with labeled initiators. The photograph of the gel shows the results of an abortive transcription initiation reaction using three different dinucleotide initiators, which were (1) ApG; (2) Biotin-ApG; and (3) 5' TAMRA-SpApG, and a terminating nucleotide, which was  $\alpha^{32}\text{P}$ -UTP. All three dinucleotides allowed for incorporation of UTP in the 3<sup>rd</sup> position to generate 5' TAMARA-SpApGpU. The unlabeled ApG incorporates more efficiently than does the Biotin-ApG, which incorporates more efficiently than the TAMARA-ApG.

[0101] FIG. 26. Abortive Transcription Initiation with a labeled terminator. The scanned image of the thin layer chromatography plate shows the results of an abortive transcription initiation reaction using an unlabeled dinucleotide initiator, ApG, and a labeled terminator, which was 5'-SF-UTP (5-thioacetemidofluorescein-uridine 5'-triphosphate. The labeled terminator was efficiently incorporated to generate the oligonucleotide product ApGpU.

[0102] FIG. 27. Portion of the contig sequence of the CDKN2A gene. The sequence represents a small portion of the contig starting at 856630 nucleotides from the start of the contig sequence. The sequence represents a CpG island. Contig number: NT\_008410.4.

[0103] FIG. 28. Schematic representation of a “capture probe” to determine the methylation status of a specific gene. Oligonucleotide probes that are specific for a region near the CpG island of the target gene are immobilized onto a microtiter plate. The DNA of interest is added to the immobilized probe and bound to the capture probe. The DNA is then chemically modified to convert unmethylated C to T, and leave methyl-C unaffected. The converted DNA can then be optionally amplified by PCR to further enhance the signal. A labeled CpG initiator is then added with an RNA polymerase and labeled nucleotide(s).

[0104] FIG. 29. Preparation of sample DNA for the CpG detection assay.

The CpG detection assay is based on the manipulation and detection of immobilized CpG islands. The CpG island strands are separated by a primer extension reaction using biotinylated primers (A). The hybrid duplexes are immobilized to a streptavidin microtiter plate (B). The sample DNA strands are modified at the 5'-ends through the removal of the terminal phosphates with phosphatase (C) and the addition of thiophosphates by polynucleotide kinase (D). The sulfur-modified strands are removed from the streptavidin plate following denaturation, and then are covalently bound to a maleimide microtiter plate (E) where one or more target site probes (TSPs) are annealed (F). It is believed that the TSPs expose the template Cs in targeted CpG sequences to potential deamination by sodium bisulfite treatment while protecting the double-stranded portions of the template-strand/TSP complexes (G). This conservation of sequence allows the replacement of the TSPs, (H), following the final denaturing in the deamination reaction. At this stage the bisulfite-treated strands are ready for the transcription reaction which will quantify the level of CpG methylation at the TSP complexes. If necessary, the deaminated strands can be amplified by PCR before adding the TSPs in H.

[0105] FIG.30 shows template sequences for the abortive transcription reactions shown in FIGS. 31-34. FIG. 30a: Poly[dG-dC] is a synthetic deoxyribonucleotide polymer of repeating dCpdG. Individual strands contain variable numbers of dinucleotide repeats. FIG 30b: Bubble complex 1 was made by annealing synthetic, partially complementary template and non-template strands. The vertical offset of the non-template strand represents the single-stranded, bubble portion of the molecule. The coordinate system is based on the downstream edge of the bubble. The unpaired bases next to the double-stranded segment are at position +1. Positions to the left (upstream) of position +1 are given negative numbers starting with -1. The coordinate system is used to indicate the position of the 3' ends of the ribonucleotide initiators. The 3' end of initiator AA is aligned at +1 and the 3' end of initiator AU is aligned at +2. The transcription reaction proceeds from left to

right from 3' end of the initiator, according to theory. FIG. 30c represents the template strand without the complementary non-template strand. The sequence is shown in the 3' to 5' orientation.

[0106] FIGS. 31a and 31b show the results of transcription of single-stranded poly[dG-dC] with *E. coli* RNA polymerase and a GpC ribonucleotide initiator.  $\gamma$ -32P GTP was the only ribonucleoside triphosphate included in the reaction. The poly[dG-dC] concentration was set to 10  $\mu$ g/25  $\mu$ l reaction. FIG. 31a represents a thin layer chromatograph of the transcription reactions. Samples (1  $\mu$ l) were spotted at the site marked GTP. The trinucleotide product GpCpG migrated from the origin while the GTP showed no mobility. FIG. 31b represents a 25% denaturing polyacrylamide gel used to electrophorese 6  $\mu$ l samples of the reactions analyzed in FIG. 31a. BPB refers to the bromophenol blue marker. Inclusion of only GTP in the reaction limited to products to the trinucleotide GpCpG.

[0107] FIG. 32 represents a thin layer chromatograph of transcription reactions employing Bubble complex 1 (samples 1-4) or the template-strand alone (samples 5-8). The reactions were carried out in the absence of Na-acetate or in the presence of 150 mM Na-acetate. Radioactive UTP was included in reactions employing ApA initiator. Radioactive ATP was added to reactions containing ApU initiator. The 3' end of ApA aligns with +1 while the 3' end of ApU is offset to +2. The trinucleotide products of ApA and ApU migrate at approximately the same rate during chromatography. All reactions contained 17 ng of bubble complex or 8.5 ng of single-strand. Samples 9 and 10 are negative controls containing radioactive ATP in high salt and low salt buffer respectively. Samples 11 and 12 are negative controls containing radioactive UTP in high- and low salt buffers.

[0108] FIG. 33 shows a thin layer chromatograph comparing transcription of Bubble complex 1 (17 ng/reaction) by different RNA polymerases. All transcription reactions were initiated with ApA. Incorporation of radioactive UTP produced the trinucleotide product ApApU. Reactions containing *E. coli* RNA polymerase Holoenzyme and Core enzyme contained 150 mM Na-

acetate. Reactions containing T7 or SP6 RNA polymerses were in low-salt reaction buffer that was not supplemented with Na-acetate. The sample labeled 'UTP' is a negative control containing radioactive UTP in high-salt transcription buffer.

[0109] FIG. 34a and 34b show thin layer chromatographs of transcription reactions performed with decreasing amounts of Bubble 1 complex. All reactions were initiated with ApA in the presence of radioactive UTP. E. coli RNA polymerase holoenzyme was present at 1.5 pmoles per reaction. The amounts of template listed in the figure were in a 25  $\mu$ l reaction volume.

[0110] FIG. 35. Incorporation of ATP analogs. Promoter T7A1 was subjected to transcription with E. coli holoenzyme, ApU initiator, CTP, UTP, radioactive GTP and various ATP analogs. ATP is first incorporated into position +5. Exclusive production of the tetramer AUCG indicates that the ATP analog is not incorporated. Extension of the tetramer to the pentamer product only, indicates that the ATP analog is a chain terminator. Production of full length transcripts indicates that the analog can be incorporated into 2 or more successive positions. The lanes represent reactions containing: 1. ATP, 2. 8-APAS-ATP, 3. Alexafluor-647-ATP, 4. BODIPY-FL-ATP, 5. Tetramethylrhodamine-ATP, 6. Texas-Red-ATP, 7. Lissamine-ATP, 8. Fluorescene-ATP, 9. Coumarin-ATP, 10. 3'-OM3-SF-ATP, 11. bis-3'-OMe - ATP, 12.  $\alpha$ -thio-ATP, 13. Tetramer standard, 14. 20-mer standard.

[0111] FIG. 36: Primer Extension of P16 DNA with SH-dUTP. 10  $\mu$ L of each sample run on an 8% polyacrylamide gel with 1X TBE buffer at 2000 V for approximately three hours. The gel was then exposed to film for twenty hours at -80°C in order to obtain this image.

[0112] FIG. 37: Primer Extension Using P16DF and SH-dUTP. The samples were heated to 85°C for 4 minutes before loading into 8% polyacrylamide gel with 0.5X TBE buffer at 2000 V for approximately three hours. The gel was then exposed to film for sixteen hours at -80°C.

[0113] FIG. 38 Primer Extension with and without SH-dUTP: The entire volume of each reaction was heated to 85°C for 4 minutes before loading into

- 59 -

8% polyacrylamide gel with 0.5X TBE buffer at 1700V for about 2.5 hours. The gel was then exposed to film for sixteen hours at -80°C to obtain the image seen in Figure 38.

[0114] FIG 39. Primers and amplication region P16DF.

[0115] FIG 40: Evaluating of nucleotide analogs as abscription substrate for *E.coli* holo RNAP during Initiation. DG142APC1 bubble complex for the promoter DG142 was the template for abscription (lanes 8-12) and evaluation of biotinylated initiators on EMAC2361-TSP2361 complex. FIG. 40A: TLC gel. autoradiogram. FIG. 40B TLC gel under UV light. FIG. 40C: PAGE gel autoradiogram. Lanes 1. ApC+ $\alpha^{32}$ P-GTP; 2. Bio-6-ApC+ $\alpha^{32}$ P-GTP; 3. ApC+GTP+ $\alpha^{32}$ P-CTP; 4. Bio-6-ApC+GTP+ $\alpha^{32}$ P-CTP; 5. Bio-6-ApC+Coum.GTP+ $\alpha^{32}$ P-CTP; 6.  $\alpha^{32}$ P-GTP control; 7.  $\alpha^{32}$ P-CTP control; 8. ApU+ATP+ $\alpha^{32}$ P-GTP; 9. ApU+Coum.ATP+ $\alpha^{32}$ P-GTP; 10. ApU+Coum.GTP+ $\alpha^{32}$ P-ATP; 11. ApU+DNP- GTP+ $\alpha^{32}$ P-ATP; 12. ApU+ $\alpha^{32}$ P-GTP; 13.  $\alpha^{32}$ P-GTP control; 14.  $\alpha^{32}$ P-ATP control; 15. Coumarin ATP control.

[0116] FIG 41. FIG 41A. Evaluation of the FRET pair made by abscription using *E.coli* holo RNA polymerase enzyme and the promoter T7A1 as the template. The reaction was initiated by PyMPO-SpApU (lanes 4-8). For comparison ApU was also used as initiator (lanes 1-3). Coumarin ATP and GTP were the analogs used (lanes 3, 6, 7 and 8) and 3'OCH<sub>3</sub>ATP was as the terminator (lanes 3 and 8). FIG 41B: Silica gel TLC of the above reaction.

[0117] FIG. 42: Autoradiogram of the TLC plate. AU-HIP/S10 complex was initiated with UpA and bis-CTP along with TCEP was added (lane 1 and 2).  $\alpha^{32}$ P-UTP was added to the reaction in lane 1. Lane 3 has SF-CTP only.

[0118] FIG. 43. Incorporation of HS-CTP during elongation by *E.coli* RNA polymerase holoenzyme. Autoradiogram of 8% polyacrylamide/7M urea is shown in (A) and that of 25% is shown in (B). Autoradiogram of TLC plate is shown in (C). RNA 20nt long complex (20 mer) was made from the T7 A1 promoter with initiator ApU (lane 1) and biotinylated ApU (lane 2 &3) and

- 60 -

CTP, ATP and  $\alpha^{32}\text{P}$ -GTP in the presence of heparin. Then it was elongated in the presence of UTP (lane 1, 2 & 3) and HS-CTP (lane 3). Samples for the TLC were treated with CIP for 1 hour at 37°C.

[0119] Figure 44: Evaluation of the analog 3'OMe-8-(2, 4-DNP-S)-ATP as a transcription substrate for E.coli RNA polymerase holoenzyme. The incorporation of the analog was analyzed by electrophoresis in a 25% polyacrylamide/7M urea gel. Transcription was initiated with ApU (Lanes 1-5, 200  $\mu\text{M}$ ) along with CTP and  $\mu$ 32P-GTP (20  $\mu\text{M}$  each). This combination of precursors produces the tetramer transcript AUCG (Lane 1). Inclusion of ATP(20  $\mu\text{M}$ ) allows the extension of the tetramer to a 20 nt long transcript (Lane 2). Substitution of the chain terminating 3'-OMe-ATP (80  $\mu\text{M}$ ) for ATP produces the pentamer AUCG-3'-OMe-A (Lane 3). Substitution of the chain terminating 3'OMe-8-(2, 4-DNP-S)-ATP analog (100  $\mu\text{M}$ , Lane 4; 500  $\mu\text{M}$ , Lane 5) for ATP produces the pentamer AUCG-3'OMe-8-(2, 4-DNP-S)-A. This pentamer has a retarded electrophoretic mobility compared to AUCG-3'-OMe-A. Lane 6 represents a negative control containing only  $\alpha^{32}\text{P}$ -GTP. Bpb and Xc represent the dye fronts for Bromophenol Blue and Xylene Cyano FF respectively.

#### DETAILED DESCRIPTION OF THE INVENTION

[0120] The invention provides, *inter alia*, methods and kits for detecting the presence of a target molecule (such as nucleic acid sequence or protein) by generating multiple detectable oligo- or polynucleotides through reiterative synthesis events on a defined nucleic acid. The methods generally comprise using a labeled nucleotide or oligonucleotide transcription initiator to initiate synthesis of an abortive oligonucleotide product that is substantially complementary to a defined site on a target nucleic acid; using a chain terminator to terminate the polymerization reaction; and, optionally, using either (1) a target site probe to form a transcription bubble complex which comprises double-stranded segments on either side of a single-stranded target

site or (2) an artificial promoter cassette comprising a transcription bubble region which includes a target site or (3) an artificial promoter cassette that is attached to any target molecule and then used to generate a signal.

[0121] In accordance with one aspect, the invention provides methods of synthesizing multiple abortive oligonucleotide transcripts from portions of a target DNA or RNA sequence, wherein the methods comprise combining and reacting the following: (a) a single-stranded target nucleic acid comprising at least one target site; (b) an RNA initiator that is complementary to a site on the target nucleic acid that is upstream of the target site; (c) an RNA polymerase; (d) optionally, nucleotides and/or nucleotide analogs; (e) a chain terminator; and (f) optionally, either (1) a target site probe that partially hybridizes to a target region on the target nucleic acid, forming a transcription bubble complex that includes first and second double-stranded regions on either side of a single-stranded target site or (2) an artificial promoter cassette comprising a transcription bubble region that includes a transcription start site. The combination is subjected to suitable conditions, as described below, such that (a) a target site probe hybridizes with a target nucleic acid in a target region that includes the target site; (b) an RNA initiator hybridizes upstream of a target site; (c) an RNA polymerase utilizes the RNA initiator to initiate transcription at the target site, elongation occurs, and an oligonucleotide transcript is synthesized; (d) a chain terminator terminates transcription during elongation; (e) the RNA polymerase releases the short, abortive oligonucleotide transcript without substantially translocating from the polymerase binding site or dissociating from the template; and (f) (c)-(e) are repeated until sufficient signal is generated and the reaction is stopped. Alternatively, (a) an artificial promoter cassette hybridizes with an end of the target nucleic acid; (b) an RNA initiator hybridizes upstream of a transcription start site; (c) an RNA polymerase utilizes the RNA initiator to initiate transcription at the target site, elongation occurs, and an oligonucleotide transcript is synthesized; (d) a chain terminator terminates transcription during elongation; (e) the RNA polymerase releases the short, abortive

oligonucleotide transcript without substantially translocating from the polymerase binding site or dissociating from the template; and (f) (c)-(e) are repeated until sufficient signal is generated and the reaction is stopped.

#### General Techniques

[0122] The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature, such as, "Molecular Biology Techniques Manual," third edition, (Coyne *et al.*, 2001); "Short Protocols in Molecular Biology," fourth edition, (Ausubel *et al.*, 1999) "Molecular Cloning: A Laboratory Manual", second edition (Sambrook *et al.*, 1989); "Oligonucleotide Synthesis" (M. J. Gait, ed., 1984); "Animal Cell Culture" (R. I. Freshney, ed., 1987); "Methods in Enzymology" (Academic Press, Inc.); "Current Protocols in Molecular Biology" (F. M. Ausubel *et al.*, eds., 1987, and periodic updates); "PCR: The Polymerase Chain Reaction" (Mullis *et al.*, eds., 1994).

[0123] Primers, initiators, oligonucleotides, and polynucleotides employed as reactants in the present invention can be generated using standard techniques known in the art or may be obtained from commercial sources, including but not restricted to Sigma/Aldrich, Molecular Probes, Trilink Technologies.

#### Terms

[0124] To facilitate understanding of the invention, the following terms have the following meanings unless expressly stated otherwise:

[0125] "About" as used herein means that a number referred to as "about" comprises the recited number plus or minus 1-10% of that recited number. For example, "about" 50 nucleotides can mean 45-55 nucleotides or as few as 49-51 nucleotides depending on the situation.

[0126] “Transcription” is an enzyme-mediated process that synthesizes a complementary RNA transcript that corresponds to a nucleic acid template sequence. Transcription typically includes three phases, namely, initiation, elongation, and termination. The transcript of the template is processively synthesized by a polymerase through the formation of a phosphodiester bond between an initiator, which may be a mononucleoside, a mononucleotide, an oligonucleotide, or polynucleotide, and a subsequent NTP, *et cetera.*, without the dissociation of either the nascent transcript or the polymerase from the template, until the polymerase reaches either a termination sequence on the template or the end of the template sequence or is stopped by other means, such as protein-DNA transcription roadblocks. As used in typical hybridization assays, the termination of transcription is generally achieved when the polymerase completes the elongation phase and reaches the end of the template sequence or a specific transcription termination signal after translocating from the initial enzyme binding site (promoter) on the template. In this context, “translocation” means that the polymerase moves along the template sequence from an initial enzyme binding site on the template to another point on the template which is at least 50 nucleotides downstream of the enzyme binding site.

[0127] “Abortive transcription” is an enzyme-mediated process that reiteratively initiates and terminates the synthesis of oligonucleotides that correspond to at least one portion, or target site, of a complementary nucleic acid template sequence. The abortive oligonucleotides synthesized vary in length of nucleotides, and may contain from about 2 to about 26 nucleotides, about 26 to about 50 nucleotides and about 50 nucleotides to about 100 nucleotides, greater than 100 nucleotides, and greater than about 500 nucleotides.

[0128] “Abortive transcription” also includes three phases, namely, initiation, elongation, and termination. During the initiation phase, a polymerase forms a phosphodiester bond between an initiator and a second NTP, and then adds subsequent NTPs, *et cetera.*, transcribing the template sequence to synthesize

an oligonucleotide transcript of from about 2 to about 50 nucleotides in length and then terminating the transcription event by releasing the nascent oligonucleotide transcript, without the polymerase substantially translocating from the polymerase binding site or dissociating from the template. In other words, the RNA polymerase substantially remains at the initial binding site on the template, releases a first nascent oligonucleotide transcript, and then is capable of engaging in another transcription initiation event to produce a second oligonucleotide transcript, which is substantially complementary to substantially the same target site that was transcribed to produce the first oligonucleotide transcript. In this manner, the polymerase reiteratively transcribes a single portion of the template (*i.e.*, a target region) and releases multiple copies of substantially identical nascent oligonucleotide transcripts.

- [0129] “Reverse transcription” refers to the transcription of an RNA template to synthesize complementary DNA (cDNA).
- [0130] “Reiterative” refers to multiple identical or highly similar copies of a sequence of interest.
- [0131] “Replication” is an enzyme-mediated process which synthesizes a complementary nucleic acid molecule from a single-stranded nucleic acid template sequence. The DNA replicate of the template is synthesized by a DNA polymerase through the formation of a phosphodiester bond between a primer and a first deoxyribonucleoside triphosphate (dNTP), followed by the formation of a second phosphodiester bond between the first dNTP and a subsequent dNTP, *et cetera.*, without the dissociation of either the DNA replicate or the DNA polymerase from the template, until the DNA polymerase reaches either a termination sequence on the template or the end of the template sequence. In a typical DNA primer extension reaction, replication of the template terminates when the DNA polymerase synthesizes the entire template sequence after translocating from the initial enzyme binding site on the template. In this context, “translocation” means that the DNA polymerase moves along the template sequence from an initial enzyme

binding site on the template to another point on the template which is downstream of the enzyme binding site.

[0132] "Oligonucleotide product" refers to the oligonucleotide that is synthesized by the reiterative synthesis reaction of the present invention. An oligonucleotide product may be an "oligonucleotide transcript," if the polymerization reaction is a transcription reaction catalyzed by an RNA polymerase, or an "oligonucleotide repeat," if the polymerization reaction is a DNA synthesis reaction catalyzed by telomerase or DNA polymerase.

[0133] "Termination" refers to the use of a chain terminator to conclude a chain elongation or primer extension reaction that is catalyzed by a polymerase. A "chain terminator" or "terminator" may comprise any compound, composition, complex, reactant, reaction condition, or process (including withholding a compound, reactant, or reaction condition) which inhibits the continuation of transcription by the polymerase beyond the initiation and/or elongation phases. A "chain terminating nucleotide" is a chain terminator that comprises a nucleotide or nucleotide analog that inhibits further chain elongation once incorporated, due to either the structure of the nucleotide analog or the sequence of the nucleic acid being copied or transcribed.

[0134] A "target sequence" or "target polynucleotide" is a polynucleotide sequence of interest for which detection, characterization or quantification is desired. The actual nucleotide sequence of the target sequence may be known or not known.

[0135] A "abortive transcription target site," when used in the context of abortive transcription, is that portion of the target sequence that is detected by transcription by a polymerase to form an oligonucleotide product. In accordance with the invention, there is at least one target site on a target nucleic acid. The sequence of a target site may or may not be known with particularity. That is, while the actual genetic sequence of the target nucleic acid may be known, the genetic sequence of a particular target site that is transcribed or replicated by a polymerase need not be known.

[0136] A “abortive transcription target region,” when used in the context of abortive transcription, is that portion of a target sequence to which a target site probe partially hybridizes to form a bubble complex, as described in detail below. In accordance with the invention, there is at least one target region on a target nucleic acid, and each target region comprises a target site. The sequence of a target region is known with sufficient particularity to permit sufficiently stringent hybridization of a complementary target site probe, such that the target site probe forms a bubble complex with the target region.

[0137] Generally, a “template” is a polynucleotide that contains the target nucleotide sequence. In some instances, the terms “target sequence”, “template polynucleotide”, “target nucleic acid”, “target polynucleotide”, “nucleic acid template”, “template sequence”, and variations thereof, are used interchangeably. Specifically, the term “template” refers to a strand of nucleic acid on which a complementary copy is synthesized from nucleotides or nucleotide analogs through the activity of a template-dependent nucleic acid polymerase. Within a duplex, the template strand is, by convention, depicted and described as the “bottom” strand. Similarly, the non-template strand is often depicted and described as the “top” strand. The “template” strand may also be referred to as the “sense” strand, and the non-template strand as the “antisense” strand.

[0138] “Synthesis” generally refers to the process of producing a nucleic acid, via chemical or enzymatic means. More typically, chemical synthesis is used for single strands of a nucleic acid. Enzyme mediated “Synthesis” encompasses both transcription and replication from a template. Synthesis includes a single copy or multiple copies of the target. “Multiple copies” means at least 2 copies. A “copy” does not necessarily mean perfect sequence complementarity or identity with the template sequence. For example, copies can include nucleotide analogs, intentional sequence alterations (such as sequence alterations introduced through a primer comprising a sequence that is hybridizable, but not complementary, to the template), and/or sequence errors that occur during synthesis.

[0139] "Polynucleotide" or "nucleic acid strand", as used interchangeably herein, refers to nucleotide polymers of any length, such as two or more, and includes both DNA and RNA. The nucleotides can be deoxyribonucleotides, ribonucleotides, nucleotide analogs (including modified phosphate moieties, bases, or sugars), or any substrate that can be incorporated into a polymer by a suitable enzyme, such as a DNA polymerase or an RNA polymerase. Thus, a polynucleotide may comprise modified nucleotides, such as methylated nucleotides, and their analogs. If present, modification to the nucleotide structure may be imparted before or after synthesis of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. Other types of modifications include, for example, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, cabamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, for example, proteins (e.g., glutathione-s-transferase, methylases, demethylases, DNA repair enzymes, nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with intercalators (e.g., ethidium, acridine, psoralen, etc.), those with antibody-specific haptens (dinitrophenyl (DNP), biotin, etc.), those with affinity tags (hexahistidine, glutathione, etc.), those containing chelators (e.g., metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, those with modified linkages (e.g., alpha anomeric nucleic acids, etc.), those with chemical or photochemical activities (DNA or RNA cleavage agents, crosslinkers, fluorescent compounds, etc.) as well as unmodified forms of the polynucleotide(s). Further, any of the hydroxyl groups ordinarily present on the pentose (*i.e.*, ribose or deoxyribose) ring of a nucleotide may be, for example, replaced by phosphonate or phosphate groups, protected by standard protecting groups, activated to prepare additional linkages to additional

nucleotides, or conjugated to a solid support. The 5' and 3' terminal OH groups on the pentose ring of a nucleotide can be phosphorylated or substituted with amines or organic capping group moieties of from about 1 to about 50 carbon atoms. Other hydroxyl groups on the ribose or deoxyribose ring may also be derivatized to standard protecting groups. Polynucleotides can also contain analogous forms of ribose or deoxyribose sugars that are generally known in the art, including, for example, 2'-O-methyl-2'-O-allyl, 2'-fluoro- or 2'-azido-ribose, carbocyclic sugar analogs, anomeric sugars, epimeric sugars, such as arabinose, xylose, pyranose sugars, furanose sugars, sedoheptuloses, acyclic analogs, and abasic nucleoside analogs such as methyl riboside. One or more phosphodiester linkages may be replaced by alternative linking groups. These alternative linking groups include, but are not limited to, embodiments wherein phosphate is replaced by P(O)S("thioate"), P(S)S ("dithioate"), "(O)NR<sub>2</sub> ("amide"), P(O)R, P(O)OR', CO or CH<sub>2</sub> ("formacetal"), in which each R or R' is independently H or substituted or unsubstituted alkyl (1-20 C) optionally containing an ether (--O--) linkage, aryl, alkenyl, cycloalkyl, cycloalkenyl, or araldyl. Not all linkages in a polynucleotide need be identical. The preceding description applies to all polynucleotides referred to herein, including RNA and DNA.

[0140] "Nucleotide" or "NTP" refers to a base-sugar-phosphate compound. "Base" refers to a nitrogen-containing ring molecule that, when combined with a pentose sugar and a phosphate group, form a nucleotide. Bases include single ring pyrimidines, such as cytosine (C), thymine (T), and uracil (U), and double ring purines, such as adenine (A) and guanine (G). "Sugar" or "pentose sugar" generally refers to a pentose ring, such as a ribose ring or deoxyribose ring. Nucleotides are the monomeric subunits of both types of nucleic acid polymers, that is, RNA and DNA. "Nucleotide" or "NTP" refers to any nucleoside 5' phosphate, that is, ribonucleoside 5' phosphates (*i.e.*, mono-, di-, and triphosphates) and deoxyribonucleoside 5' phosphates (*i.e.*, mono-, di-, and triphosphates), and includes "nucleoside phosphate analogs", "nucleotide analogs", and "NTP analogs". "Nucleoside phosphate analog",

“nucleotide analog”, and “NTP analog” refer to any nucleoside 5’ phosphate (*i.e.*, mono-, di-, or triphosphate) which is analogous to a native nucleotide but which contains one or more chemical modifications when compared to the corresponding native nucleotide. Nucleotide analogs include base-modified analogs. Particularly included are 5-S-pyrimidines and 8-S-purines, and modified versions thereof, as described elsewhere herein, and may comprise modified forms of deoxyribonucleotides as well as ribonucleotides.

[0141] “Nucleoside” refers to a base-sugar combination without a phosphate group. Nucleosides include, but are not limited to, adenosine (A), cytidine (C), guanosine (G), thymidine (T), and uridine (U).

[0142] The term “oligonucleotide” generally refers to short, typically single-stranded, synthetic polynucleotides that may also include “mononucleotides” containing one nucleotide and “dinucleotides” containing two nucleotides and polynucleotides that are less than about 200 nucleotides in length, from about 200 to about 500 nucleotides, or about 500 or more nucleotides in length. More typically, an oligonucleotide may be defined as a molecule comprised of two or more nucleotides, including deoxyribonucleotides and/or ribonucleotides. The exact size depends on many factors, which in turn depend on the ultimate function or use of the oligonucleotide. The oligonucleotide may be generated in any manner, including chemical synthesis, DNA replication, degradation of longer DNA or RNA, transcription, reverse transcription, abortive transcription or reiterative synthesis, as further described herein, and a combination thereof.

[0143] Because mononucleotides undergo a reaction which synthesizes oligonucleotides by covalently bonding the 3’ oxygen of a first mononucleotide pentose ring to the 5’ phosphate of a second mononucleotide through a phosphodiester linkage, a first end of an oligonucleotide is referred to as the “5’ end” if the 5’ phosphate of the terminal nucleotide is not linked to a 3’ oxygen of a nucleotide pentose ring, and a second end of an oligonucleotide is referred to as the “3’ end” if the 3’ oxygen of the terminal nucleotide is not linked to a 5’ phosphate of a subsequent nucleotide pentose

ring. As used herein, a nucleic acid sequence, even if the sequence is internal to a larger oligonucleotide, also may be said to have 5' and 3' ends. For single-stranded DNA or RNA, a first region along a nucleic acid strand is said to be "upstream" of a second region, if the 3' end of the first region is before the 5' end of the second region when moving along a strand of nucleic acid in a 5' → 3' direction. Conversely, a first region along a nucleic acid strand is said to be "downstream" of a second region, if the 5' end of the first region is after the 3' end of the second region when moving along a strand of nucleic acid in a 5' → 3' direction.

[0144] The term "3'" generally refers to a region or position in a polynucleotide or oligonucleotide that is 3' (downstream) from another region or position in the same polynucleotide or oligonucleotide when moving along the polynucleotide or oligonucleotide in a 5' → 3' direction.

[0145] The term "5'" generally refers to a region or position in a polynucleotide or oligonucleotide that is 5' (upstream) from another region or position in the same polynucleotide or oligonucleotide when moving along the polynucleotide or oligonucleotide in a 5' → 3' direction.

[0146] "Nucleic acid sequence" refers to an oligonucleotide or polynucleotide, and fragments, segments, or portions thereof, and to DNA or RNA of genomic or synthetic origin, which may be single- or double-stranded, and represents either the sense or the antisense strand.

[0147] The term "substantially single-stranded", when used in reference to a nucleic acid substrate, means that the substrate molecule exists primarily as a single strand of nucleic acid in contrast to a double-stranded substrate which exists as two substantially complementary segments or regions of nucleic acid that are held together by inter-strand or intra-strand base pairing interactions.

[0148] As used herein, the terms "complementary" or "complementarity" are used in reference to a first polynucleotide (which may be an oligonucleotide) which is in "antiparallel association" with a second polynucleotide (which also may be an oligonucleotide). As used herein, the term "antiparallel association" refers to the alignment of two polynucleotides such that

individual nucleotides or bases of the two associated polynucleotides are paired substantially in accordance with Watson-Crick base-pairing rules. For example, the sequence "A-G-T" is complementary to the sequence "T-C-A." Complementarity may be "partial," in which only some of the polynucleotides' bases are matched according to the base pairing rules. Or, there may be "complete" or "total" complementarity between the polynucleotides. The degree of complementarity between the polynucleotides has significant effects on the efficiency and strength of the hybridization between two polynucleotides. This is of particular importance in synthesis reactions, as well as detection methods which depend upon binding between polynucleotides. Those skilled in the art of nucleic acid technology can determine duplex stability empirically by considering a number of variables, including, for example, the length of the first polynucleotide, which may be an oligonucleotide, the base composition and sequence of the first polynucleotide, and the ionic strength and incidence of mismatched base pairs. A general formula that may be used to calculate the melting temperature of an oligonucleotide is:  $T_m = (2(UA) + 4(GC)) - 0.5C$  for every 1% formamide. For DNA-DNA hybrids, the  $T_m$  is approximated by the following formula:  $T_m = 81.5 + 16.6 (\log M) + 0.41 (\%G + C) - 500/L$ ;  $M$  is the molarity of the monovalent cations;  $L$  is the length of the hybrid base pairs (Anal Biochem. 138:267-284, 1984).

[0149] The terms "self-complementary" and "self-complementarity", when used in reference to a polynucleotide (e.g., an oligonucleotide), mean that separate regions of the polynucleotide can base-pair with each other. Because this term refers only to intramolecular base-pairing, any strand said to have a region of self-complementarity must have at least two regions capable of base-pairing with one another. As defined above, complementarity may be either "complete" or "partial". As used in reference to the oligonucleotides of the present invention, regions of an oligonucleotide are considered to have significant self-complementarity when these regions are capable of forming a duplex of at least 3 contiguous base pairs (*i.e.*, three base pairs of complete

complementarity), or when they may form a longer duplex that is partially complementary.

[0150] The term "primer" generally refers to a short, single-stranded oligonucleotide which has a free 3'-OH group and which can bind to and hybridize with a target sequence that is potentially present in a sample of interest. After hybridizing to a target sequence, a primer is capable of promoting or initiating polymerization or synthesis of a polynucleotide or oligonucleotide extension product that is complementary to the target sequence or a portion of the target sequence. A primer is selected to be "substantially" complementary to a specific portion of a target nucleic acid sequence. A primer is sufficiently complementary to hybridize with a target sequence and facilitate either transcription or replication of a portion of the target nucleic acid. A primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment may be attached to the 5' end of the primer, with the remainder of the primer sequence being substantially complementary to the template strand. Non-complementary bases can be interspersed within the primer, provided that the primer sequence has sufficient complementarity with the template sequence to hybridize with the template and thereby form a template-primer complex for initiating synthesis of a polynucleotide or oligonucleotide product.

[0151] The terms "purine" and "pyrimidine" are accorded their customary and usual meaning as understood by those of ordinary skill in the art. Purines comprise a six-membered and a five-membered nitrogen-containing ring, fused together. Purines include 6 amino purine (adenine), 2-amino-6-oxy purine (guanine), 6-oxy purine (hypoxanthine) and 2,6,dioxypurine (xanthine). Pyrimidines have only a six-membered nitrogen-containing ring. Pyrimidines include 2,4-dioxy pyrimidine (uracil), 2,4-dioxy-5-methyl pyrimidine (thymine), 2-oxy-4-amino pyrimidine (Cytosine) and 2,4-dioxy-6-carboxy pyrimidine (Orotic acid). As used in the present invention, the terms purine and pyrimidine encompass both ribonucleotides, deoxyribonucleotides

and dideoxyribonucleotides, unphosphorylated and mono, di and triphosphorylated forms, and both free and incorporated into a nucleic acid.

[0152] The term "initiator" refers to a mononucleoside, mononucleotide, oligonucleotide, polynucleotide or analog thereof, which is incorporated into the 5' end of a nascent RNA molecule and may be considered a "primer" for RNA synthesis ("initiator primer").

[0153] In one embodiment, an RNA initiator facilitates the initiation of transcription at a target site on a single-stranded target nucleic acid in the absence of a template promoter sequence, as is known in the art. (See, U.S. Pat. No. 5,571,669; Daube and von Hippel, *Science*, 258: 1320-1324 (1992)). In another embodiment, initiators are used to randomly start abortive transcription at a plurality of target sites on the nucleic acid template (FIG. 14). The initiators and/or the individual nucleotides or nucleotide analogs that are used to extend the initiators may be suitably modified to enable signal generation, detection of the oligonucleotide products, and a determination of the presence or absence of the target sequence.

[0154] For example, it may be desirable to modify the initiator to provide the initiator with a label moiety for a variety of purposes, including detection of an abortive oligonucleotide product(s). Examples of such modifications are described elsewhere herein.

[0155] As used herein, the term "elongator" is a mononucleoside or mononucleotide, which is incorporated via the elongator's 5' end to the 3' end of the preceding nucleotide. The 3' of the newly-incorporated nucleotide may then be used for the attachment to a yet further nucleotide. An elongator nucleotide may function as an initiator or terminator under certain reaction conditions.

[0156] As used herein, the term "hybridization" is used in reference to the base-pairing of complementary nucleic acids, including polynucleotides and oligonucleotides. Hybridization and the strength of hybridization (*i.e.*, the strength of the association between the nucleic acids) is impacted by such factors as the degree of complementarity between the nucleic acids, the

stringency of the reaction conditions involved, the melting temperature ( $T_m$ ) of the formed hybrid, and the G:C ratio within the duplex nucleic acid. Generally, "hybridization" methods involve annealing a complementary polynucleotide to a target nucleic acid (*i.e.*, the sequence to be detected either by direct or indirect means). The ability of two polynucleotides and/or oligonucleotides containing complementary sequences to locate each other and anneal to one another through base pairing interactions is a well-recognized phenomenon.

[0157] With regard to complementarity, it may be important for some diagnostic applications to determine whether the hybridization of two polynucleotides and/or oligonucleotides represents complete or partial complementarity. For example, where it is desired to detect simply the presence or absence of a pathogen (such as from a virus, bacterium, fungi, mycoplasma, or protozoan for example), the hybridization method need only ensure that hybridization occurs when the relevant sequence is present. Conditions can be selected where both partially complementary probes and completely complementary probes will hybridize. Other diagnostic applications, however, may require that the hybridization method be capable of distinguishing between partial and complete complementarity, such as in cases where it may be of interest to detect a genetic polymorphism, that is, a difference in a single base pair between multiple alleles (variations) that may exist for a particular gene or genetic marker.

[0158] "Stringency" generally refers to the conditions under which nucleic acid hybridizations are conducted, including temperature, ionic strength, and the presence of other compounds. Conditions of "high stringency" generally refer to those conditions under which nucleic acid base pairing will occur only between polynucleotide and/or oligonucleotide regions that have a high frequency of complementary base sequences. Consequently, conditions of "weak" or "low" stringency may be preferred when it is desirable to hybridize or anneal two polynucleotides and/or oligonucleotides, which are not completely complementary to one another.

[0159] The term "reactant" is used in its broadest sense. A reactant can comprise an enzymatic reactant, a chemical reactant, or ultraviolet light (ultraviolet light, particularly short wavelength ultraviolet light, is known to break polynucleotide polymers). Any agent capable of reacting with an oligonucleotide or polynucleotide to modify the oligonucleotide or polynucleotide is encompassed by the term "reactant," including a "reactant nucleotide" that is added to a reaction mixture for incorporation into an oligonucleotide product by a polymerase.

[0160] A "complex" is an assembly of components. A complex may or may not be stable and may be directly or indirectly detected. For example, as described herein, given certain components of a reaction and the type of product(s) of the reaction, the existence of a complex can be inferred. For example, in the abortive transcription method described herein, a complex is generally an intermediate with respect to a final reiterative synthesis product, such as a final abortive transcription or replication product for example.

[0161] A "reaction mixture" is an assemblage of components, which, under suitable conditions, react to form a complex (which may be an intermediate) and/or a product(s).

[0162] The term "enzyme binding site" refers to a polynucleotide region that is characterized by a sequence and/or structure that is capable of binding to a particular enzyme or class of enzymes, such as a polymerase.

[0163] "Polymerase" refers to any agent capable of facilitating or catalyzing the polymerization (joining) of nucleotides and/or nucleotide analogs. Suitable agents include naturally occurring enzymes, such as naturally occurring RNA polymerases (including RNA-dependent and DNA-dependent RNA polymerases), DNA polymerases (including DNA-dependent and RNA-dependent DNA polymerases), as well as modified or mutant enzymes that may currently exist (such as the mutant RNA polymerases disclosed in Sousa, et al., U.S. Pat. No. 6,107,037 for example) or may be hereafter created or designed, which modified or mutant enzymes may be designed to exhibit characteristics that are desirable for particular applications. Exemplary

characteristics of a modified or mutant enzyme may include, but are not limited to, relaxed template specificity, relaxed substrate specificity, increased thermostability, and/or the like. It is intended that the term "polymerase" encompasses both thermostable and thermolabile enzymes.

[0164] The term "thermostable" when used in reference to an enzyme, such as an RNA or DNA polymerase for example, indicates that the enzyme is functional or active (*i.e.*, can perform catalysis) at an elevated temperature, typically at about 55 °C or higher. Thus, a thermostable polymerase can perform catalysis over a broad range of temperatures, including temperatures both above and below about 55 °C.

[0165] A "terminal transferase" is an enzyme that catalyses the template independent addition of deoxynucleotide phosphates to the 3'-end of a nucleic acid.

[0166] The term "template-dependent polymerase" refers to a nucleic acid polymerase that synthesizes a polynucleotide or oligonucleotide product by copying or transcribing a template nucleic acid, as described above, and which does not synthesize a polynucleotide in the absence of a template. This is in contrast to the activity of a template-independent nucleic acid polymerase, such as terminal deoxynucleotidyl transferase or poly-A polymerase for example, that may synthesize or extend nucleic acids in the absence of a template.

[0167] A "DNA-dependent RNA polymerase" is an enzyme which facilitates or catalyzes the polymerization of RNA from a complementary DNA template.

[0168] A "DNA-dependent DNA polymerase" is an enzyme which facilitates or catalyzes DNA replication or synthesis, that is, the polymerization of DNA from a complementary DNA template.

[0169] An "RNA-dependent RNA polymerase" is an enzyme which facilitates or catalyzes the polymerization of RNA from a complementary RNA template.

[0170] An "RNA-dependent DNA polymerase" or "reverse transcriptase" is an enzyme that facilitates or catalyzes the polymerization of DNA from a complementary RNA template.

[0171] "Primer extension", "extension", "elongation", and "extension reaction" is the sequential addition of nucleotides to the 3' hydroxyl end of a mononucleotide, oligonucleotide, or polynucleotide initiator or primer which has been annealed or hybridized to a longer, template polynucleotide, wherein the addition is directed by the nucleic acid sequence of the template and/or the binding position of the polymerase. Extension generally is facilitated by an enzyme capable of synthesizing a polynucleotide or oligonucleotide product from a primer or initiator, nucleotides and a template. Suitable enzymes for these purposes include, but are not limited to, any of the polymerases described above.

[0172] "Incorporation" refers to becoming a part of a nucleic acid polymer. There is a known flexibility in the terminology regarding incorporation of nucleic acid precursors. For example, the nucleotide dGTP is a deoxyribonucleoside triphosphate. Upon incorporation into DNA, dGTP becomes dGMP, that is, a deoxyguanosine monophosphate moiety. Although DNA does not include dGTP molecules, one may say that one incorporates dGTP into DNA.

[0173] The terms "sample" and "test sample" are used in their broadest sense. For example, a "sample" or "test sample" is meant to include a specimen or culture (*e.g.*, microbiological cultures) as well as both biological and environmental samples. Samples of nucleic acid used in the methods of the invention may be aqueous solutions of nucleic acid derived from a biological or environmental sample and separated, by methods known in the art, from other materials, such as proteins, lipids, and the like, that may be present in the sample and that may interfere with the methods of the invention or significantly increase the "background" signal in carrying out the methods.

[0174] A biological sample may comprise any substance which may include nucleic acid, such as animal (including human) tissue, animal fluids (such as

blood, saliva, mucosal secretions, semen, urine, sera, cerebral or spinal fluid, pleural fluid, lymph, sputum, fluid from breast lavage, and the like), animal solids (e.g., stool), cultures of microorganisms, liquid and solid food and feed products, waste, cosmetics, or water that may be contaminated with a microorganism, or the like. An environmental sample may include environmental material, such as surface matter, soil, water, and industrial samples, as well as samples obtained from food and dairy processing instruments, apparatus, equipment, utensils, and disposable and non-disposable items. These examples are merely illustrative and are not intended to limit the sample types applicable to the present invention.

[0175] "Purified" or "substantially purified" refers to nucleic acids that are removed from their natural environment, isolated or separated, and are at least 25% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated. An "isolated polynucleotide" or "isolated oligonucleotide" is therefore a substantially purified polynucleotide or oligonucleotide.

[0176] The term "gene" refers to a DNA sequence that comprises control and coding sequences necessary for the production of a polypeptide or precursor. The polypeptide can be encoded by a full-length coding sequence or by any portion of the coding sequence, so long as the desired functional activity is retained.

[0177] A "deletion" is defined as a change in a nucleic acid sequence in which one or more nucleotides are absent as compared to a standard nucleic acid sequence.

[0178] An "insertion" or "addition" is a change in a nucleic acid sequence which has resulted in the addition of one or more nucleotides as compared to a standard nucleic acid sequence.

[0179] A "substitution" results from the replacement of one or more nucleotides in a nucleic acid by different nucleotides.

[0180] An "alteration" in a nucleic acid sequence refers to any change in a nucleic acid sequence or structure, including, but not limited to a deletion, an

addition, an addition-deletion, a substitution, an insertion, a reversion, a transversion, a point mutation, or a microsatellite alteration, or methylation.

[0181] “Methylation” refers to the addition of a methyl group (-CH<sub>3</sub>) to a nucleotide base in DNA or RNA.

[0182] Sequence “mutation” refers to any sequence alteration in a sequence of interest in comparison to a reference sequence. A reference sequence can be a wild type sequence or a sequence to which one wishes to compare a sequence of interest. A sequence mutation includes single nucleotide changes, or alterations of more than one nucleotide in a sequence, due to mechanisms such as substitution, deletion, or insertion. A single nucleotide polymorphism (SNP) is also a sequence mutation as used herein.

[0183] “Microarray” and “array,” as used interchangeably herein, refer to an arrangement of a collection of polynucleotide sequences in a centralized location. Arrays can be on a solid substrate, such as a glass slide, or on a semi-solid substrate, such as nitrocellulose membrane. The polynucleotide sequences can be DNA, RNA, or any combinations thereof.

[0184] The term “label” refers to any atom, molecule, or moiety which can be used to provide a detectable (preferably quantifiable) signal, either directly or indirectly, and which can be attached to a nucleotide, nucleotide analog, nucleoside mono-, di-, or triphosphate, nucleoside mono-, di-, or triphosphate analog, polynucleotide, or oligonucleotide. Labels may provide signals that are detectable by fluorescence, radioactivity, chemiluminescence, electrical, paramagnetism, colorimetry, gravimetry, X-ray diffraction or absorption, magnetism, enzymatic activity, and the like. A label may be a charged moiety (positive or negative charge) or, alternatively, may be charge neutral. One aspect of the invention is directed to the use of particular labelled nucleotides.

[0185] “Detection” includes any means of detecting, including direct and indirect detection. Direct detection includes methods such as detection of radioactive decay, or fluorescent tags etc. Indirect detection typically requires the use of a second molecule to identify the molecule of interest, such as the use of streptavidin to detect the presence of biotin labelled nucleic acids.

[0186] "Detectably fewer" may be observed directly or indirectly, and the term indicates any reduction in the number of products (including no products). Similarly, "detectably more" products means any increase, whether observed directly or indirectly.

[0187] As used herein, the terms "comprises," "comprising", "includes", and "including", or any other variations thereof, are intended to cover a non-exclusive inclusion, such that a process, method, composition, reaction mixture, kit, or apparatus that comprises a list of elements does not include only those elements but may include other elements not expressly listed or inherent to such process, method, composition, reaction mixture, kit, or apparatus.

[0188] "A," "an," "the," and the like, unless otherwise indicated, include plural forms.

[0189] The term "alkyl" as employed herein by itself or as part of another group refers to both straight and branched chain radicals of up to 10 carbons, unless the chain length is otherwise limited, such as methyl, ethyl, propyl, isopropyl, butyl, *t*-butyl, isobutyl, pentyl, hexyl, isohexyl, heptyl, 4,4-dimethylpentyl, octyl, 2,2,4-trimethylpentyl, nonyl, or decyl.

[0190] The term "alkenyl" is used herein to mean a straight or branched chain radical of 2-10 carbon atoms, unless the chain length is otherwise limited, wherein there is at least one double bond between two of the carbon atoms in the chain, including, but not limited to, ethenyl, 1-propenyl, 2-propenyl, 2-methyl-1-propenyl, 1-butenyl, 2-butenyl, and the like. Preferably, the alkenyl chain is 2 to 8 carbon atoms in length, most preferably from 2 to 4 carbon atoms in length.

[0191] The term "alkynyl" is used herein to mean a straight or branched chain radical of 2-10 carbon atoms, unless the chain length is otherwise limited, wherein there is at least one triple bond between two of the carbon atoms in the chain, including, but not limited to, ethynyl, 1-propynyl, 2-propynyl, and the like. Preferably, the alkynyl chain is 2 to 8 carbon atoms in length, most preferably from 2 to 4 carbon atoms in length.

[0192] In all instances herein where there is an alkenyl or alkynyl moiety as a substituent group, the unsaturated linkage, *i.e.*, the vinyl or ethenyl linkage, is preferably not directly attached to a nitrogen, oxygen or sulfur moiety.

[0193] The term "alkoxy" or "alkyloxy" refers to any of the above alkyl groups linked to an oxygen atom. Typical examples are methoxy, ethoxy, isopropoxy, *sec*-butyloxy, and *t*-butyloxy.

[0194] The term "aryl" as employed herein by itself or as part of another group refers to monocyclic or bicyclic aromatic groups containing from 6 to 12 carbons in the ring portion, preferably 6-10 carbons in the ring portion. Typical examples include phenyl, biphenyl, naphthyl or tetrahydronaphthyl.

[0195] The term "aralkyl" or "arylalkyl" as employed herein by itself or as part of another group refers to alkyl groups as discussed above having an aryl substituent, such as benzyl, phenylethyl or 2-naphthylmethyl.

[0196] The term "heteroaryl" as employed herein refers to groups having 5 to 14 ring atoms; 6, 10 or 14 pi electrons shared in a cyclic array; and containing carbon atoms and 1, 2, 3, or 4 oxygen, nitrogen or sulfur heteroatoms (where examples of heteroaryl groups are: thienyl, benzo[b]thienyl, naphtho[2,3-b]thienyl, thianthrenyl, furyl, pyranyl, isobenzofuranyl, benzoxazolyl, chromenyl, xanthenyl, phenoxathiinyl, 2*H*-pyrrolyl, pyrrolyl, imidazolyl, pyrazolyl, pyridyl, pyrazinyl, pyrimidinyl, pyridazinyl, indolizinyl, isoindolyl, 3*H*-indolyl, indolyl, indazolyl, purinyl, 4*H*-quinolizinyl, isoquinolyl, quinolyl, phthalazinyl, naphthyridinyl, quinazolinyl, cinnolinyl, pteridinyl, 4*αH*-carbazolyl, carbazolyl,  $\beta$ -carbolinyl, phenanthridinyl, acridinyl, perimidinyl, phenanthrolinyl, phenazinyl, isothiazolyl, phenothiazinyl, isoxazolyl, furazanyl, phenoxazinyl, and tetrazolyl groups).

[0197] The phrase "saturated or partially unsaturated heterocycle" as employed herein, by itself or as part of another group, refers to a saturated or partially unsaturated ring system having 5 to 14 ring atoms selected from carbon atoms and 1, 2, 3, or 4 oxygen, nitrogen, or sulfur heteroatoms. Typical saturated examples include pyrrolidinyl, imidazolidinyl, pyrazolidinyl, tetrahydrofuranyl, tetrahydropyranyl, piperidyl, piperazinyl, quinuclidinyl,

morpholinyl, and dioxacyclohexyl. Typical partially unsaturated examples include pyrrolinyl, imidazolinyl, pyrazolinyl, dihydropyridinyl, tetrahydropyridinyl, and dihydropyranyl. Either of these systems can be optionally fused to a benzene ring.

[0198] The terms "heteroarylalkyl" or "heteroaralkyl" as employed herein both refer to a heteroaryl group attached to an alkyl group. Typical examples include 2-(3-pyridyl)ethyl, 3-(2-furyl)-*n*-propyl, 3-(3-thienyl)-*n*-propyl, and 4-(1-isoquinolinyl)-*n*-butyl.

[0199] The term "cycloalkyl" as employed herein by itself or as part of another group refers to cycloalkyl groups containing 3 to 9 carbon atoms. Typical examples are cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, cyclooctyl and cyclononyl.

[0200] The term "cycloalkylalkyl" or "cycloalkyl(alkyl)" as employed herein, by itself or as part of another group, refers to a cycloalkyl group attached to an alkyl group. Typical examples are 2-cyclopentylethyl, cyclohexylmethyl, cyclopentylmethyl, 3-cyclohexyl-*n*-propyl, and 5-cyclobutyl-*n*-pentyl.

[0201] The term "cycloalkenyl" as employed herein, by itself or as part of another group, refers to cycloalkenyl groups containing 3 to 9 carbon atoms and 1 to 3 carbon-carbon double bonds. Typical examples include cyclopropenyl, cyclobutenyl, cyclopentenyl, cyclohexenyl, cyclohexadienyl, cycloheptenyl, cycloheptadienyl, cyclooctenyl, cyclooctadienyl, cyclooctatrienyl, cyclononenyl, and cyclonadienyl.

[0202] The term "halogen" or "halo" as employed herein by itself or as part of another group refers to chlorine, bromine, fluorine or iodine.

[0203] The term "monoalkylamine" or "monoalkylamino" as employed herein by itself or as part of another group refers to the group NH<sub>2</sub> wherein one hydrogen has been replaced by an alkyl group, as defined above.

[0204] The term "dialkylamine" or "dialkylamino" as employed herein by itself or as part of another group refers to the group NH<sub>2</sub> wherein both hydrogens have been replaced by alkyl groups, as defined above.

[0205] The term "hydroxyalkyl" as employed herein refers to any of the above alkyl groups wherein one or more hydrogens thereof are substituted by one or more hydroxyl moieties.

[0206] The term "haloalkyl" as employed herein refers to any of the above alkyl groups wherein one or more hydrogens thereof are substituted by one or more halo moieties. Typical examples include fluoromethyl, difluoromethyl, trifluoromethyl, trichloroethyl, trifluoroethyl, fluoropropyl, and bromobutyl.

[0207] The term "carboxyalkyl" as employed herein refers to any of the above alkyl groups wherein one or more hydrogens thereof are substituted by one or more carboxylic acid moieties.

[0208] The term "heteroatom" is used herein to mean an oxygen atom ("O"), a sulfur atom ("S") or a nitrogen atom ("N"). It will be recognized that when the heteroatom is nitrogen, it may form an  $NR^aR^b$  moiety, wherein  $R^a$  and  $R^b$  are, independently from one another, hydrogen or  $C_1$  to  $C_8$  alkyl, or together with the nitrogen to which they are bound form a saturated or unsaturated 5-, 6- or 7-membered ring.

Nucleotides and nucleotide analogs.

[0209] In one aspect, the invention describes nucleotide analogs, and the use of such analogs in molecules, nucleic acids, methods, assays and kits. The term nucleotide would be understood by those of skill in the art to include both unmodified nucleotides and nucleotide analogs. Such may be labeled or unlabeled; may be suitable for incorporation into RNA or DNA; and may be modified to permit further extension or cause termination of synthesis. The choice of nucleotide will depend on the use to which it is put. One of ordinary skill in the art may determine which nucleotide or nucleotide analogue may be appropriate for any given purpose.

[0210] The detection and identification of mono, di, oligo- or polynucleotide products are facilitated by label moieties on the primer, initiator and/or on the NTPs or NTP analogs that are incorporated into each mon, di, oligo or

polynucleotide product that is synthesized on the target nucleic acid. The label or reporter moieties may be chemically or enzymatically incorporated into the nucleotides forming the primer and/or into the reactant NTPs or NTP analogs that are utilized by the polymerase during the extension reaction, or other molecules, and may include, for example, fluorescent tags; paramagnetic groups; chemiluminescent groups; metal binding sites; intercalators; photochemical crosslinkers; antibody-specific haptens; metals; small molecules which are members of a specific binding pair (such as biotin and streptavidin for example); and any other reporter moiety or moieties which can produce a detectable and/or quantifiable signal either directly or indirectly. Exemplary nucleotide analogs may include, for example, 8-S-modified purines (8-APAS-ATP) (Costas, Hanna, et al., Nucleic Acids Research 28: 1849-58 (2000)); 5-modified pyrimidines (5-APAS-UTP; 5-APAS-CTP) (Hanna, M., Meth Enzymology 180: 383-409 (1989); Hanna, M., Nucleic Acids Research 21: 2073-79 (1993)); fluorescent ribonucleotides (5-SF-UTP) (Hanna, M. et al., Nucleic Acid Research 27: 1369-76 (1999)); and haptен-tagged deoxynucleotide precursors (5-DNP-SdU) (Meyer and Hanna, Bioconjugate Chem 7: 401-412 (1996); U.S. Pat. Nos. 6,008,334 and 6,107,039).

[0211] In one embodiment, a fluorophore moiety is attached to the 5' end of the initiator/primer that is used to initiate transcription/prime replication of the target nucleic acid. In another embodiment, a fluorophore moiety is attached to the 5 or 8 position of the base of an NTP or NTP analog that is used by the polymerase to extend from an initiator or primer. In a further embodiment, a first fluorophore moiety is attached to the initiator/primer and a second fluorophore is attached to an NTP or NTP analog that is used to extend the initiator/primer. In this latter embodiment, a fluorescent energy transfer mechanism can be used, wherein the first fluorophore (e.g. fluorescein, aedans, coumarin, etc.) is excited and the emission is read from the second fluorophore (e.g. fluorescein, AEDANS, coumarin, etc.) when the second fluorophore is brought into proximity with the first fluorophore by the polymerase during synthesis of the oligonucleotide/polynucleotide product.

Alternatively, the first and second fluorophores may function by an electron transfer mechanism, wherein the first fluorophore absorbs energy from the second fluorophore when the polymerase brings the first and second fluorophores into proximity with each other, and the first fluorophore releases the energy in a radiative manner, thereby enabling detection.

[0212] In one aspect, a first fluorophore is a fluorescent energy donor, which is attached to a first reactant (*i.e.*, either a nucleotide that is incorporated into the initiator/primer or a nucleotide that is to be incorporated by the polymerase into the oligonucleotide/polynucleotide product), and a second fluorophore is a fluorescent energy acceptor, which is attached to a second reactant (either a nucleotide that is incorporated into the initiator/primer nucleotide or a nucleotide that is to be incorporated by the polymerase into the oligonucleotide/polynucleotide product) that is different from the first reactant. In one embodiment, each of the four NTPs or NTP analogs that may be used to extend the primer/initiator is tagged with a unique fluorescent energy acceptor which is capable of a distinct emission wavelength when brought into proximity with the fluorescent energy donor on the primer. Preferably, the fluorescent energy transfer can be measured in real time, without isolation of the oligonucleotide/polynucleotide products, since neither the initiator nor unincorporated NTPs or NTP analogs alone will produce a signal at the wavelength used for detection.

[0213] Fluorescent and chromogenic molecules and their relevant optical properties are amply described in the literature. See, for example, Berlman, *Handbook of Fluorescence Spectra of Aromatic Molecules*, 2nd Edition (Academic Press, New York, 1971); Griffiths, *Colour and Constitution of Organic Molecules* (Academic Press, New York, 1976); Bishop, ed., *Indicators* (Pergamon Press, Oxford, 1972); Haugland, *Handbook of Fluorescent Probes and Research Chemicals* (Molecular Probes, Eugene, 1992); Pringsheim, *Fluorescence and Phosphorescence* (Interscience Publishers, New York, 1949); and the like. Further, there is guidance in the literature for derivatizing fluorophore and quencher molecules for covalent

attachment via common reactive groups that can be added to a nucleotide, as exemplified by the following references: Haugland (supra); Ullman et al., U.S. Pat. No. 3,996,345; Khanna et al., U.S. Pat. No. 4,351,760; Costas, Hanna, et al., Nucleic Acids Research 28: 1849-58 (2000); Hanna, M. et al., Nucleic Acid Research 27: 1369-76 (1999); and Meyer and Hanna, Bioconjugate Chem 7: 401-412 (1996).

[0214] In general, nucleotide labeling can be accomplished through any of a large number of known nucleotide labeling techniques using known linkages, linking groups, and associated complementary functionalities. Suitable donor and acceptor moieties that can effect fluorescence resonance energy transfer (FRET) include, but are not limited to, 4-acetamido-4'-isothiocyanatostilbene-2,2'-disulfonic acid; acridine and derivatives: acridine, acridine isothiocyanate; 5-(2'-aminoethyl)amminonaphthalene-1-sulfonic acid (EDANS); 4-amino-N-[3-vinylsulfonyl]phenyl]naphthalimide-3,5 disulfonate; N-(4-amino-1-naphthyl)maleimide; anthranilamide; BODIPY; Brilliant Yellow; coumarin, and derivatives: coumarin, 7-amino-4-methylcoumarin (AMC, Coumarin 120), 7-amino-4-trifluoromethylcouluarin (Coumaran 151); cyanine dyes; cyanosine; 4',6-diaminidino-2-phenylindole (DAPI); 5', 5"-dibromopyrogallol-sulfonaphthalein (Bromopyrogallol Red); 7-diethylamino-3-(4'-isothiocyanatophenyl)-4-methylcoumarin; diethylenetriamine pentaacetate; 4,4'-diisothiocyanatodihydro-stilbene-2,2'-disulfonic acid; 4,4'-diisothiocyanatostilbene-2,2'-disulfonic acid; 5-[dimethylamino]naphthalene-1-sulfonyl chloride (DNS, dansylchloride); 4-dimethylaminophenylazophenyl-4'-isothiocyanate (DABITC); eosin and derivatives: eosin, eosin isothiocyanate; erythrosin and derivatives: erythrosin B, erythrosin, isothiocyanate; ethidium; fluorescein and derivatives: 5-carboxyfluorescein (FAM), 5-(4,6-dichlorotriazin-2-yl)aminofluorescein (DTAF), 2',7'-dimethoxy-4'5'-dichloro-6-carboxyfluorescein (JOE), fluorescein, fluorescein isothiocyanate, QFITC, (XRITC); fluorescamine; IR144; IR1446; Malachite Green isothiocyanate; 4-methylumbelliferoneortho cresolphthalein; nitrotyrosine; pararosaniline; Phenol Red; B-phycoerythrin; o-

phthaldialdehyde; pyrene and derivatives: pyrene, pyrene butyrate, succinimidyl 1pyrene; butyrate quantum dots; Reactive Red 4 (Cibacron™ Brilliant Red 3B-A); rhodamine and derivatives: 6-carboxy-X-rhodamine (ROX), 6-carboxyrhodamine (R6G), lissamine rhodamine B, sulfonyl chloride rhodamine (Rhod), rhodamine B, rhodamine 123, rhodamine X isothiocyanate, sulforhodamine B, sulforhodamine 101, sulfonyl chloride derivative of sulforhodamine 101 (Texas Red); N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA); tetramethyl rhodamine; tetramethyl rhodamine isothiocyanate (TRITC); riboflavin; rosolic acid; terbiun chelate derivatives; Cy 3; Cy 5; Cy 5.5; Cy 7; IRD 700; IRD 800; La Jolla Blue; phthalo cyanine; and naphthalo cyanine.

[0215] There are many linking moieties and methodologies for attaching fluorophores to nucleotides, as exemplified by the following references: Eckstein, ed., Oligonucleotides and Analogues: A Practical Approach (IRL Press, Oxford, 1991); Zuckerman et al., Nucleic Acids Research 15: 5305-5321(1987) (3' thiol group on oligonucleotide); Sharma et al., Nucleic Acids Research 19: 3019 (1991) (3' sulfhydryl); Giusti et al., PCR Methods and Applications 2: 223-227 (1993); Fung et al., U.S. Pat. No. 4,757,141 (5' phosphoamino group via Aminolink™ II, available from Applied Biosystems, Foster City, Calif.); Stabinsky, U.S. Pat. No. 4,739,044 (3' aminoalkylphosphoryl group); Agrawal et al., Tetrahedron Letters 31: 1543-1546 (1990) (attachment via phosphoramidate linkages); Sproat et al., Nucleic Acids Research 15: 4837 (1987) (5-mercaptop group); Nelson et al., Nucleic Acids Research 17: 7187-7194 (1989) (3' amino group); Hanna, M., Meth Enzymology 180: 383-409 (1989); Hanna, M., Nucleic Acids Research 21: 2073-79 (1993); Hanna, M. et al., Nucleic Acid Research 27: 1369-76 (1999) (5-mercaptop group); Costas, Hanna, et al., Nucleic Acids Research 28: 1849-58 (2000) (8-mercaptop group); and the like.

[0216] In accordance with the invention, detection of the oligonucleotide products is indicative of the presence of a target sequence. Quantitative analysis is also feasible. Direct and indirect detection methods (including

quantitation) are well known in the art. For example, by comparing the amount of oligonucleotide/polynucleotide products that are generated from a test sample containing an unknown amount of a target nucleic acid to an amount of oligonucleotide products that were generated from a reference sample that has a known quantity of a target nucleic acid, the amount of a target nucleic acid in the test sample can be determined.

[0217] Preferred nucleotide analogs for use in the present invention include, 8-modified purines; 5-modified pyrimidines; fluorescent ribonucleotides; and haptens-tagged deoxynucleotide precursors. Such analogs are suitable for use in DNA or RNA molecules comprising the compound; in methods of use of the analogs, including those involving labelling, synthesizing, or detecting the presence of a nucleic acid, protein or other item of interest; and in kits comprising the compound.

[0218] Suitable nucleotide analogs for use in the method/kit/molecules of the invention comprise 8-S-substituted purines and 5-S-substituted pyrimidines of the formula NucSR;

wherein Nuc is pyrimidinyl or purinyl;

wherein S is sulfur;

[0219] and wherein R is selected from the group consisting of H, haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, a C<sub>6-10</sub> aryl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> arylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylcarbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonylamino(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyl(C<sub>1-6</sub>)alkyl and C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-4 substituents independently selected from the group consisting of halo,

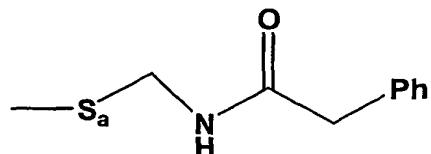
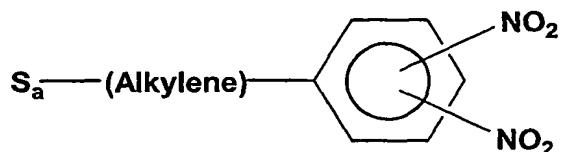
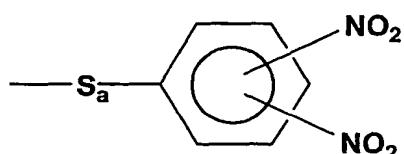
hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, C<sub>6-10</sub> arylamino, (C<sub>6-10</sub> aryl)carbonylamino, mono(C<sub>6-10</sub> aryl)aminocarbonyl, di(C<sub>6-10</sub> aryl)aminocarbonyl, mono(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, di(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, N-(C<sub>6-10</sub>)aryl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>6-10</sub> aryl)aminocarbonyl, C<sub>1-6</sub> alkylthio, C<sub>6-10</sub> arylthio, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro, cyano, heteroaryl and saturated or partially unsaturated heterocycle, wherein the heteroaryl and saturated or partially unsaturated heterocycle are independently monocyclic or fused bicyclic and independently have 5 to 10 ring atoms, wherein one or more of the ring atoms are independently selected from the group consisting of oxygen, nitrogen and sulfur; *provided that* (1) when the aryl portion of R is substituted with two substituents, at least one of the substituents is other than nitro, and (2) when R is phenylacetamidomethyl, the phenyl portion is substituted at least once;

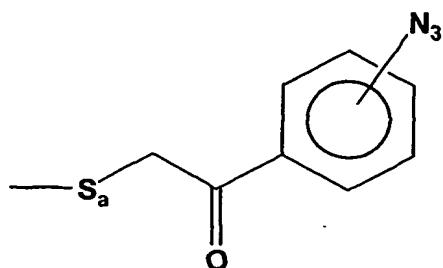
[0220] In one embodiment, R is selected from the group consisting of H, haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinkers, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, phenyl, phenyl(C<sub>1-6</sub>)alkyl, phenylamino(C<sub>1-6</sub>)alkyl, phenoxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>alkyl)carbonylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>alkyl)carbonyloxy(C<sub>1-6</sub>)alkyl, benzoylcarbonylamino(C<sub>1-6</sub>)alkyl, benzoyloxy(C<sub>1-6</sub>)alkyl, benzoyl(C<sub>1-6</sub>)alkyl and phenyl(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-2 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub>

- 90 -

haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>) alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, phenylamino, benzoylamino, phenylaminocarbonyl, diphenylaminocarbonyl, phenyl(C<sub>1-6</sub> alkyl) aminocarbonyl, di(phenyl(C<sub>1-6</sub> alkyl))aminocarbonyl, N-phenyl-N-(C<sub>1-6</sub> alkyl) aminocarbonyl, N-phenyl(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-phenyl (C<sub>1-6</sub>)alkyl-N-phenylaminocarbonyl, C<sub>1-6</sub> alkylthio, phenylthio, phenyl(C<sub>1-6</sub>) alkylthio, carboxy(C<sub>1-6</sub>)alkyl, nitro and cyano.

[0221] In another embodiment the SR is





wherein a is 1 or 2.

[0222] In one embodiment R is Coumerin. In one embodiment R is Tetramethylrhodamine. In one embodiment R is Lissamine. In one embodiment R is Alexafluor. In one embodiment R is BODIPY-FL. In one embodiment R is Fluorescene. In one embodiment R is 3'-OMe-SF. In one embodiment R is Texas-Red. In one embodiment R is bis-3'-OMe-ATP. In one embodiment R is 3'-OMe-SH-ATP. In one embodiment R is TAMARA. In one embodiment R is Biotin. In one embodiment R is H. In one embodiment R is PyMPO.

[0223] In some embodiments, SR is substituted with S-L-R, wherein L is a linker having a backbone chain length of 1 to about 8 atoms; wherein the atoms are selected from the group consisting of carbon, sulfur, nitrogen, oxygen and phosphorous.

[0224] In some embodiments, R is selected from the group consisting of haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, and a nucleic acid cellular uptake group,

[0225] In some embodiments, R is selected from the group consisting of H, phenyl, phenyl(C<sub>1-6</sub>)alkyl, phenylamino(C<sub>1-6</sub>)alkyl, phenoxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub> alkyl)carbonylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub> alkyl)carbonyloxy(C<sub>1-6</sub>)

alkyl, benzoylcarbonylamino(C<sub>1-6</sub>)alkyl, benzyloxy(C<sub>1-6</sub>)alkyl, benzoyl(C<sub>1-6</sub>)alkyl and phenyl(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-2 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, phenylamino, benzoylamino, phenylaminocarbonyl, diphenylaminocarbonyl, phenyl(C<sub>1-6</sub> alkyl)aminocarbonyl, di(phenyl(C<sub>1-6</sub> alkyl))aminocarbonyl, N-phenyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-phenyl(C<sub>1-6</sub> alkyl)-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-phenyl(C<sub>1-6</sub>)alkyl-N-phenylaminocarbonyl, C<sub>1-6</sub> alkylthio, phenylthio, phenyl(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro and cyano.

[0226] In further embodiments R is selected from the group consisting of phenyl, phenyl(C<sub>1-4</sub>)alkyl, phenyl(C<sub>1-4</sub> alkyl)carbonylamino (C<sub>1-4</sub>)alkyl and benzoyl(C<sub>1-4</sub>)alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-2 substituents independently selected from the group consisting of halo, hydroxyl, C<sub>1-4</sub> alkyl, C<sub>1-4</sub> haloalkyl, C<sub>1-4</sub> hydroxyalkyl, C<sub>1-4</sub> alkoxy, (C<sub>1-4</sub> alkyl)carbonyl, (C<sub>1-4</sub> alkoxy)carbonyl, amino, amino(C<sub>1-4</sub>)alkyl, aminocarbonyl, mono(C<sub>1-4</sub> alkyl)aminocarbonyl, di(C<sub>1-4</sub> alkyl)aminocarbonyl, C<sub>1-4</sub> alkylamino, di(C<sub>1-4</sub>)alkylamino, (C<sub>1-4</sub> alkyl)carbonylamino, phenylamino, benzoylamino, phenylaminocarbonyl, phenyl(C<sub>1-4</sub> alkyl)aminocarbonyl, C<sub>1-4</sub> alkylthio, phenylthio, phenyl(C<sub>1-4</sub>)alkylthio, alkylthio, carboxy, carboxy(C<sub>1-4</sub>)alkyl, nitro and cyano.

[0227] When SR is a protected thiol, R may be useful in the present invention when R includes C<sub>6-10</sub> aryl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> arylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> aryloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonylamino(C<sub>1-6</sub>)alkyl, C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl)carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonylamino(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)

carbonyloxy(C<sub>1-6</sub>)alkyl, (C<sub>6-10</sub> aryl)carbonyl(C<sub>1-6</sub>)alkyl and C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl) carbonyl(C<sub>1-6</sub>)alkyl, each of which is optionally substituted.

[0228] More preferably, when SR is a protected thiol, R includes phenyl, phenyl(C<sub>1-6</sub>)alkyl, phenylamino(C<sub>1-6</sub>)alkyl, phenoxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>) alkylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub>)alkyloxy(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub> alkyl) carbonylamino(C<sub>1-6</sub>)alkyl, phenyl(C<sub>1-6</sub> alkyl)carbonyloxy(C<sub>1-6</sub>)alkyl, benzoylcarbonylamino(C<sub>1-6</sub>)alkyl, benzoyloxy(C<sub>1-6</sub>)alkyl, benzoyl(C<sub>1-6</sub>)alkyl and phenyl(C<sub>1-6</sub> alkyl)carbonyl(C<sub>1-6</sub>)alkyl, each of which is optionally substituted.

[0229] Other useful values of R include phenyl, phenyl(C<sub>1-4</sub>)alkyl, phenyl(C<sub>1-4</sub> alkyl)carbonylamino(C<sub>1-4</sub>)alkyl and benzoyl(C<sub>1-4</sub>)alkyl, each of which is optionally substituted.

[0230] Useful values of R include optionally-substituted phenyl groups.

[0231] Useful values of R also include optionally-substituted phenyl(C<sub>1-4</sub>)alkyl groups. Especially useful phenyl(C<sub>1-4</sub>)alkyl groups include phenylmethyl, 2-phenylethyl, 3-phenylpropyl and 4-phenylbutyl, each of which is optionally substituted.

[0232] Useful values of R also include optionally-substituted benzoyl(C<sub>1-4</sub>)alkyl groups. Especially useful benzoyl(C<sub>1-4</sub>)alkyl groups include benzoylmethyl, 2-benzoylethyl, 3-benzoylpropyl and 4-benzoylbutyl, each of which is optionally substituted.

[0233] Particularly useful values of R also include optionally-substituted phenyl(C<sub>1-4</sub> alkyl)carbonylamino(C<sub>1-4</sub>)alkyl groups. Especially useful phenyl(C<sub>1-4</sub> alkyl)carbonylamino(C<sub>1-4</sub>)alkyl groups include phenylacetamidomethyl, 2-(phenylacetamido)-ethyl, 3-(phenylacetamido)-propyl, 4-(phenylacetamido)-butyl, 3-phenylpropanoylaminomethyl, 4-phenylbutanoylaminomethyl and 2-(3-phenylpropanoylamino)-ethyl, each of which is optionally substituted.

[0234] The aryl portion of R is optionally substituted 1-4 times. When the aryl portion of R is substituted, it is preferably substituted 1 or 2 times.

[0235] When the aryl portion of R is substituted, useful substituents include halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, C<sub>6-10</sub> arylamino, (C<sub>6-10</sub> aryl)carbonylamino, mono(C<sub>6-10</sub> aryl)aminocarbonyl, di(C<sub>6-10</sub> aryl)aminocarbonyl, mono(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, di(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, N-(C<sub>6-10</sub>)aryl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>6-10</sub> aryl)aminocarbonyl, C<sub>1-6</sub> alkylthio, C<sub>6-10</sub> arylthio, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro, cyano, heteroaryl and saturated or partially unsaturated heterocycle, wherein the heteroaryl and saturated or partially unsaturated heterocycle are independently monocyclic or fused bicyclic and independently have 5 to 10 ring atoms, wherein one or more of the ring atoms are independently selected from the group consisting of oxygen, nitrogen and sulfur.

[0236] More useful substituents on the aryl portion of R include halo, hydroxyl, C<sub>1-6</sub> alkyl, C<sub>3-8</sub> cycloalkyl, C<sub>1-6</sub> haloalkyl, C<sub>1-6</sub> hydroxyalkyl, C<sub>2-6</sub> alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, phenylamino, benzoylamino, phenylaminocarbonyl, diphenylaminocarbonyl, phenyl(C<sub>1-6</sub> alkyl)aminocarbonyl, di(phenyl(C<sub>1-6</sub> alkyl))aminocarbonyl, N-phenyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-phenyl(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-phenyl(C<sub>1-6</sub>)alkyl-N-phenylaminocarbonyl, C<sub>1-6</sub> alkylthio, phenylthio, phenyl(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro and cyano.

[0237] Other useful substituents on the aryl portion of R include halo, hydroxyl, C<sub>1-4</sub> alkyl, C<sub>1-4</sub> haloalkyl, C<sub>1-4</sub> hydroxyalkyl, C<sub>1-4</sub> alkoxy, (C<sub>1-4</sub> alkyl)carbonyl, (C<sub>1-4</sub> alkoxy)carbonyl, amino, amino(C<sub>1-4</sub>)alkyl, aminocarbonyl,

mono(C<sub>1-4</sub> alkyl)aminocarbonyl, di(C<sub>1-4</sub> alkyl)aminocarbonyl, C<sub>1-4</sub> alkylamino, di(C<sub>1-4</sub>)alkylamino, (C<sub>1-4</sub> alkyl)carbonylamino, phenylamino, benzoylamino, phenylaminocarbonyl, phenyl(C<sub>1-4</sub> alkyl)aminocarbonyl, C<sub>1-4</sub> alkylthio, phenylthio, phenyl(C<sub>1-4</sub>)alkylthio, carboxy, carboxy(C<sub>1-4</sub>)alkyl, nitro and cyano.

[0238] Particularly useful substituents on the aryl portion of R include halo groups. Especially useful halo groups include chloro and bromo. Particularly useful substituents on the aryl portion of R include hydroxyl groups. Particularly useful substituents on the aryl portion of R also include C<sub>1-4</sub> alkyl groups. Especially useful C<sub>1-4</sub> alkyl groups include methyl, ethyl, *n*-propyl, isopropyl and *n*-butyl. Particularly useful substituents on the aryl portion of R also include C<sub>1-4</sub> haloalkyl groups. Especially useful C<sub>1-4</sub> haloalkyl groups include methyl substituted 1-3 times with a halogen atom independently selected from fluoro, chloro and bromo; and ethyl substituted 1-3 times with a halogen atom independently selected from fluoro, chloro and bromo.

[0239] Particularly useful substituents on the aryl portion of R also include C<sub>1-4</sub> hydroxyalkyl groups. Especially useful C<sub>1-4</sub> hydroxyalkyl groups include hydroxymethyl, hydroxyethyl, hydroxypropyl and dihydroxypropyl. Particularly useful substituents on the aryl portion of R also include C<sub>1-4</sub> alkoxy groups. Especially useful C<sub>1-4</sub> alkoxy groups include methoxy, ethoxy, *n*-propoxy, isopropoxy and *n*-butoxy. Particularly useful substituents on the aryl portion of R also include (C<sub>1-4</sub> alkyl)carbonyl groups. Especially useful (C<sub>1-4</sub> alkyl)carbonyl groups include acetyl, propanoyl, 2-methylpropanoyl and butanoyl. Particularly useful substituents on the aryl portion of R also include (C<sub>1-4</sub> alkoxy)carbonyl groups. Especially useful (C<sub>1-4</sub> alkoxy)carbonyl groups include methoxycarbonyl, ethoxycarbonyl, *n*-propoxycarbonyl, isopropoxycarbonyl and *n*-butoxycarbonyl.

[0240] Particularly useful substituents on the aryl portion of R also include amino groups. Particularly useful substituents on the aryl portion of R also include amino(C<sub>1-4</sub>)alkyl groups. Especially useful amino(C<sub>1-4</sub>)alkyl groups include aminomethyl, aminoethyl, amino-*n*-propyl, aminoisopropyl and

amino-*n*-butyl. Particularly useful substituents on the aryl portion of R also include aminocarbonyl groups.

[0241] Particularly useful substituents on the aryl portion of R also include mono(C<sub>1-4</sub> alkyl)aminocarbonyl groups. Especially useful mono(C<sub>1-4</sub> alkyl)aminocarbonyl groups include methylaminocarbonyl, ethylaminocarbonyl, *n*-propylaminocarbonyl, isopropylaminocarbonyl and *n*-butyl aminocarbonyl.

[0242] Particularly useful substituents on the aryl portion of R also include di(C<sub>1-4</sub> alkyl)aminocarbonyl groups. Especially useful di(C<sub>1-4</sub> alkyl)aminocarbonyl groups include dimethylaminocarbonyl, diethylaminocarbonyl, *N*-methyl-*N*-ethylaminocarbonyl, *N*-methyl-*N*-*n*-propylaminocarbonyl and *N*-methyl-*N*-*n*-butylaminocarbonyl. Particularly useful substituents on the aryl portion of R also include C<sub>1-4</sub> alkylamino groups. Especially useful C<sub>1-4</sub> alkylamino groups include methylamino, ethylamino, *n*-propylamino, isopropylamino and *n*-butylamino. Particularly useful substituents on the aryl portion of R also include di(C<sub>1-4</sub>)alkylamino groups. Especially useful di(C<sub>1-4</sub>)alkylamino groups include dimethylamino, diethylamino, *N*-methyl-*N*-ethylamino, *N*-methyl-*N*-*n*-propylamino and *N*-methyl-*N*-*n*-butylamino.

[0243] Particularly useful substituents on the aryl portion of R also include (C<sub>1-4</sub> alkyl)carbonylamino groups. Especially useful (C<sub>1-4</sub> alkyl)carbonylamino groups include acetylamino, propanoylamino, 2-methylpropanoylamino and butanoylamino. Particularly useful substituents on the aryl portion of R also include phenylamino. Particularly useful substituents on the aryl portion of R also include benzoylamino. Particularly useful substituents on the aryl portion of R also include phenylaminocarbonyl. Particularly useful substituents on the aryl portion of R also include phenyl(C<sub>1-4</sub> alkyl)aminocarbonyl groups. Especially useful phenyl(C<sub>1-4</sub> alkyl)aminocarbonyl groups include phenylmethylaminocarbonyl, 2-phenylethylaminocarbonyl, 3-phenylpropyl aminocarbonyl and 4-phenylbutylaminocarbonyl. Particularly useful substituents on the aryl portion of R also include C<sub>1-4</sub> alkylthio groups.

Especially useful  $C_{1-4}$  alkylthio groups include methylthio, ethylthio, *n*-propylthio, isopropylthio and *n*-butylthio.

[0244] Particularly useful substituents on the aryl portion of R also include phenylthio. Particularly useful substituents on the aryl portion of R also include phenyl( $C_{1-4}$ )alkylthio groups. Especially useful phenyl( $C_{1-4}$ )alkylthio groups include phenylmethylthio, 2-phenylethylthio, 3-phenylpropylthio and 4-phenylbutylthio. Particularly useful substituents on the aryl portion of R also include carboxy. Particularly useful substituents on the aryl portion of R also include carboxy( $C_{1-4}$ )alkyl groups. Especially useful carboxy( $C_{1-4}$ )alkyl groups include carboxymethyl, 2-carboxyethyl, 3-carboxypropyl and 4-carboxybutyl. Particularly useful substituents on the aryl portion of R also include nitro. Particularly useful substituents on the aryl portion of R also include cyano.

[0245] In one embodiment, R is selected from the group consisting of fluorescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS and dinitrophenol, pyrene, stilbene, coumarine, bimane, naphthalene, pyridoxazole, naphthalimid, NBD, and BODIPY.

[0246] In one embodiment, R is selected from the group consisting of haptens, biotin, an enzyme (e.g. horseradish peroxidase or alkaline phosphatase), a protein, an artificial promoter cassette, a photocrosslinker, a chemical crosslinker, streptavidin, a fluorescent moiety, a colorimetric moiety, a luminescent moiety, a chemiluminescent moiety, a metal (e.g. gold, silver), a dye, a nucleic acid cellular uptake group, a  $C_{6-10}$  aryl,  $C_{6-10}$  ar( $C_{1-6}$ )alkyl,  $C_{6-10}$  arylamino( $C_{1-6}$ )alkyl,  $C_{6-10}$  aryloxy( $C_{1-6}$ )alkyl,  $C_{6-10}$  ar( $C_{1-6}$ )alkylamino( $C_{1-6}$ )alkyl,  $C_{6-10}$  ar( $C_{1-6}$ )alkyloxy( $C_{1-6}$ )alkyl,  $C_{6-10}$  ar( $C_{1-6}$ )alkylcarbonylamino( $C_{1-6}$ )alkyl,  $C_{6-10}$  ar( $C_{1-6}$  alkyl)carbonyloxy( $C_{1-6}$ )alkyl, ( $C_{6-10}$  aryl)carbonylamino( $C_{1-6}$ )alkyl, ( $C_{6-10}$  aryl)carbonyloxy( $C_{1-6}$ )alkyl, ( $C_{6-10}$  aryl)carbonyl( $C_{1-6}$ )alkyl and  $C_{6-10}$  ar( $C_{1-6}$  alkyl)carbonyl( $C_{1-6}$ )alkyl, wherein the aryl portion of each of the preceding groups is optionally substituted with 1-4 substituents independently selected from the group consisting of halo, hydroxyl,  $C_{1-6}$  alkyl,  $C_{3-8}$  cycloalkyl,  $C_{1-6}$  haloalkyl,  $C_{1-6}$  hydroxyalkyl,  $C_{2-6}$

alkenyl, C<sub>2-6</sub> alkynyl, C<sub>1-6</sub> alkoxy, (C<sub>1-6</sub> alkyl)carbonyl, (C<sub>1-6</sub> alkoxy)carbonyl, amino, amino(C<sub>1-6</sub>)alkyl, aminocarbonyl, mono(C<sub>1-6</sub> alkyl)aminocarbonyl, di(C<sub>1-6</sub> alkyl)aminocarbonyl, C<sub>1-6</sub> alkylamino, di(C<sub>1-6</sub>)alkylamino, (C<sub>1-6</sub> alkyl)carbonylamino, C<sub>6-10</sub> arylamino, (C<sub>6-10</sub> aryl)carbonylamino, mono(C<sub>6-10</sub> aryl)aminocarbonyl, di(C<sub>6-10</sub> aryl)aminocarbonyl, mono(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, di(C<sub>6-10</sub> ar(C<sub>1-6</sub> alkyl))aminocarbonyl, N-(C<sub>6-10</sub>)aryl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>1-6</sub> alkyl)aminocarbonyl, N-(C<sub>6-10</sub>)ar(C<sub>1-6</sub>)alkyl-N-(C<sub>6-10</sub> aryl)aminocarbonyl, C<sub>1-6</sub> alkylthio, C<sub>6-10</sub> arylthio, C<sub>6-10</sub> ar(C<sub>1-6</sub>)alkylthio, carboxy, carboxy(C<sub>1-6</sub>)alkyl, nitro, cyano, heteroaryl and saturated or partially unsaturated heterocycle, wherein the heteroaryl and saturated or partially unsaturated heterocycle are independently monocyclic or fused bicyclic and independently have 5 to 10 ring atoms, wherein one or more of the ring atoms are independently selected from the group consisting of oxygen, nitrogen and sulfur; *provided that* (1) when the aryl portion of R is substituted with two substituents, at least one of the substituents is other than nitro, and (2) when R is phenylacetamidomethyl, the phenyl portion is substituted at least once;

[0247] and R is not selected from the group consisting of florescein, APAS, H, TAMRA, EADANS, EADANS, IAEDANS and dinitrophenol, pyrene, stilbene, coumarine, bimane, naphthalene, pyridoxazole, naphthalamide, NBD, and BODIPY.

[0248] Some species of the above nucleotide analogs have been previously described, including, 8-modified purines (8-APAS-ATP) (Costas, Hanna, et al., Nucleic Acids Research 28: 1849-58 (2000)); 5-modified pyrimidines (5-APAS-UTP; 5-APAS-CTP) (Hanna, M., Meth Enzymology 180: 383-409 (1989); Hanna, M., Nucleic Acids Research 21: 2073-79 (1993)); fluorescent ribonucleotides (5-SF-UTP) (Hanna, M. et al., Nucleic Acid Research 27: 1369-76 (1999)); and hapten-tagged deoxynucleotide precursors (5-DNP-SdU) (Meyer and Hanna, Bioconjugate Chem 7: 401-412 (1996); U.S. Pat. Nos. 6,008,334 and 6,107,039). All citations are incorporated herein by reference in their entirities.

### Labelled Molecules

[0249] In one embodiment of the invention, nucleic acids molecules comprise at least one nucleotide analog. Nucleic acids comprising such analogs may be DNA, RNA or synthetic variants of such molecules, and may comprise one, two or more analogs, and each nucleic acid molecule may have more than one type of analog. Nucleic acids containing such analogs are useful in the practice of the methods, assays and kits of the invention. The presence of such analogs are useful in the conduct of methods, assays or kits of the invention, and may be used to enable detection of nucleic acids containing such analogs.

### Incorporation of nucleotide analogs into nucleic acids

[0250] Nucleotide analogs may be incorporated into nucleic acids by means known in the art, both chemical and enzymatic. These include polymerase chain reaction; nick translation; reverse transcription; transcription; primer extension; terminal transferase additions; ligation; end labelling; and abortive transcription. Examples of enzyme mediated systems include those using DNA-dependant DNA polymerase, such as PCR, nick translation labelling, in-vitro replication, and reactions used in DNA sequencing. Also included are DNA-dependant RNA polymerases (transcription), which includes transcription and abortive transcription; RNA-dependant DNA polymerase (reverse transcription); and RNA-dependant RNA polymerases, as mediated by various viral polymerases.

### Polymerase

[0251] Template-dependent polymerases for use in the methods, compositions and kits of the present invention are known in the art. Either eukaryotic or prokaryotic polymerases may be used. In one embodiment, the template-dependent polymerase is a thermostable polymerase. In another embodiment, the polymerase is able to tolerate label moieties on the phosphate group, the nucleic acid, and/or on the pentose ring of unincorporated nucleotides. In one

embodiment, the polymerase is a DNA-dependent RNA polymerase which is capable of transcribing a single-stranded DNA template without a promoter sequence. In another embodiment, the polymerase is a DNA-dependent RNA polymerase which is capable of transcribing a single-stranded DNA template having a promoter sequence that is capable of binding the particular RNA polymerase being used. In another embodiment, the polymerase is a DNA-dependent DNA polymerase that is capable of replicating a DNA target site to form a DNA oligonucleotide product. In a further embodiment, the polymerase is an RNA-dependent DNA polymerase that is capable of synthesizing a single-stranded complementary DNA transcript from an RNA template. Examples of suitable polymerases include the RNA polymerases encoded by *Escherichia coli*, *Thermus aquaticus*, *Escherichia coli* bacteriophage T7, *Escherichia coli* bacteriophage T3, and *Salmonella typhimurium* bacteriophage SP6; RNA-dependent RNA polymerases, such as poliovirus RNA polymerase; reverse transcriptases, such as HIV reverse transcriptase; and DNA polymerases such as *Escherichia coli*, T7, T4 DNA polymerase, Taq thermostable DNA polymerase, terminal transferase, primase, and telomerase.

[0252] In general, the enzymes included in the methods of the present invention preferably do not produce substantial degradation of the nucleic acid components produced by the methods.

#### Labeling and Detection

[0253] In accordance with the invention, the polymerase catalyzes a reaction in the usual 5'→ 3' direction on the oligonucleotide product and either transcribes or replicates the target nucleic acid by extending the 3' end of the initiator or primer through the sequential addition of nucleotides (NTPs), which may include nucleotide analogs (NTP analogs) and which may be labeled or unlabeled.

[0254] In another aspect, labeling may be achieved through the use of terminal transferase addition to the 3'-OH end, or the use of ligase, to add a nucleotide to the 3' end. In another aspect, labeling is achieved via chemical coupling.

[0255] In accordance with an aspect of the invention, detectable oligonucleotide or polynucleotide products are synthesized from a target nucleic acid template. The detection and identification of oligonucleotide or polynucleotide products is facilitated by label moieties on the initiator and/or on the NTPs or NTP analogs that are incorporated by the polymerase into each oligonucleotide product that is synthesized on the target nucleic acid and/or on other molecules which are part of the synthetic complex or which interact with one or more components of the synthetic complex. The label or reporter moieties may be chemically or enzymatically incorporated into the nucleotides forming the primer and/or into the reactant NTPs or NTP analogs that are utilized by the polymerase during the extension reaction.

[0256] Substituted purines and pyrimidines may also be labelled with one or more radio-isotopes, including, but not limited  $^{35}\text{S}$ ,  $^3\text{H}$  and  $^{32}\text{P}$ .

#### Target Nucleic Acid

[0257] The target nucleic acid can be either a naturally occurring or synthetic polynucleotide segment, and it can be obtained or synthesized by techniques that are well-known in the art. A target sequence to be detected in a test sample may be present initially as a discrete molecule, so that the sequence to be detected constitutes the entire nucleic acid, or may be present as only one component of a larger molecule. The target nucleic acid can be only a minor fraction of a complex mixture, such as a biological sample, and can be obtained from various biological materials by procedures that are well-known in the art. The target nucleic acid to be detected may include nucleic acids from any source, in purified, or unpurified form, which can be DNA (including double-stranded (ds) DNA and single-stranded (ss) DNA) or RNA (including tRNA, mRNA, rRNA), mitochondrial DNA or RNA, chloroplast

DNA or RNA, DNA-RNA hybrids, or mixtures thereof; genes, chromosomes, or plasmids; and the genomes of biological material, such as the genomes of microorganisms (including bacteria, yeast, viruses, viroids, molds, and fungi), plants, animals, humans, or fragments thereof. Standard techniques in the art are used to obtain and purify the nucleic acids from a test sample. Methods for the extraction and/or purification of such nucleic acids have been described, for example, by Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual (New York, Cold Spring Harbor Laboratory, third edition, 2000). Detection of an RNA target may or may not require initial complementary DNA (cDNA) synthesis, as known in the art. Detection of a DNA-RNA hybrid may require denaturation of the hybrid to obtain a ssDNA or denaturation followed by reverse transcription to obtain a cDNA.

#### Target Proteins

[0258] In another embodiment of the invention, the target may be another molecule, such as a protein, which is labeled by covalent or noncovalent attachment of a defined nucleic acid sequence which can be used as a target for further detection via PCR, hybridization, or reiterative oligonucleotide synthesis (FIG. 21). The target protein can be either a naturally occurring or synthetic polypeptide segment, and it can be obtained or synthesized by techniques that are well-known in the art. A target protein to be detected in a test sample may be present initially as a discrete molecule, so that the protein to be detected constitutes the entire protein, or may be present as only one component of a larger complex. The target protein can be only a minor fraction of a complex mixture, such as a biological sample, and can be obtained from various biological materials by procedures that are well-known in the art. The target protein to be detected may include proteins from any source, in purified or unpurified form. Standard techniques in the art are used to obtain and purify the proteins from a test sample. Methods for the extraction and/or purification of such proteins have been described, for example, by

Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual (New York, Cold Spring Harbor Laboratory, third edition, 2000).

### Immobilization

[0259] In one embodiment of the invention, the target molecule may be immobilized. In another embodiment, the target molecule may be immobilized to form, for example, a microarray. A single molecule array in accordance with this embodiment includes a solid matrix, a bioreactive or bioadhesive layer, and a bioresistant layer. Solid phases that are useful as a matrix for the present invention include, but are not limited to, polystyrene, polyethylene, polypropylene, polycarbonate, or any solid plastic material in the shape of test tubes, beads microparticles, dip-sticks, or the like. Additionally, matrices include, but are not limited to, membranes, microtiter plates (e.g., 96-well and 384-well), test tubes, and Eppendorf tubes. Solid phases also include glass beads, glass test tubes, and any other appropriate shape that is made of glass. A functionalized solid phase, such as plastic or glass, which has been modified so that the surface carries carboxyl, amino, hydrazide, or aldehyde groups can also be used. In general, suitable solid matrices comprise any surface to which a bioadhesive layer, such as a ligand-binding agent, can be attached or any surface which itself provides a ligand attachment site.

[0260] The bioadhesive layer can be an ionic adsorbent material such as gold, nickel, or copper (Montemagno and Bachand, Constructing Nanomechanical Devices Powered by Biomolecular Motors, *Nanotechnology*, 10: 225-231 (1999)), protein-adsorbing plastics, such as polystyrene (U.S. Pat. No. 5,858,801), or a covalent reactant, such as a thiol group. To create a patterned array in the bioadhesive layer, an electron-sensitive polymer, such as polymethyl methacrylate (PMMA) for example, can be used to coat the solid support and can be etched in any desired pattern with an electron beam followed by development to remove the sensitized polymer. The etched

portions of the polymer are then coated with a metal, such as nickel, and the polymer is removed with a solvent, leaving a pattern of metal posts on the substrate. This method of electron beam lithography provides the high spatial resolution and small feature size, which facilitates the immobilization of a single molecule at each point in the patterned array. An alternate means for creating high-resolution patterned arrays is atomic force microscopy. A further means is X-ray lithography.

[0261] Antibody or oligonucleotide capture probes can be attached to the bioadhesive pattern by providing a tag (such as a polyhistidine tag) on the capture probe that binds to the bioadhesive patterns (such as nickel). The capture probes may be, for example, from about 15 to about 500 nucleotides in length. Other conventional means for attachment employ homobifunctional and heterobifunctional crosslinking reagents. Homobifunctional reagents carry two identical functional groups, whereas heterobifunctional reagents contain two dissimilar functional groups to link the capture probes to the bioadhesive. The heterobifunctional cross-linking agents may contain a primary amine-reactive group and a thiol-reactive group. Covalent crosslinking agents are selected from reagents capable of forming disulfide (S—S), glycol (—CH(OH)—CH(OH)—), azo (—N=N—), sulfone (—S(=O<sub>2</sub>)—), ester (—C(=O)—O—), or amide (—C(=O)—N—) bridges. Crosslinking agents include, but are not limited to, maleamides, iodoacetamides, and disulfides. Table 1 provides a list of representative classes of crosslinking reagents and their group specificity (Wong, S.S. Chemistry of Protein Conjugation and Cross-Linking, 1991, CRC Press, Inc., Boca Raton, USA).

[0262] A bioresistant layer may be placed or superimposed upon the bioadhesive layer either before or after attachment of the capture probe to the bioadhesive layer. The bioresistant layer is any material that does not bind the capture probe. Non-limiting examples include bovine serum albumin, gelatin, lysozyme, octoxynol, polysorbate 20 (polyethenesorbitan monolaurate), and polyethylene oxide containing block copolymers and surfactants (U.S. Pat.

- 105 -

No. 5,858,801). Deposition of the bioadhesive and bioresistant layers may be accomplished by conventional means, including spraying, immersion, and evaporative deposition (metals).

Table 1: Crosslinking Reagents and group specificity

Reagent	Group specificity
alpha-haloacetyl compounds eg ICH <sub>2</sub> COOH	SH, S-CH <sub>3</sub> , NH <sub>2</sub> , phenolic, imadazole
N-maleimides	SH, NH <sub>2</sub>
mercurials	SH
Disulfides	SH
Aryl halides	SH, NH <sub>2</sub> , phenolic, imidazole
Acid anhydrides eg. Succinic anhydride	NH <sub>2</sub> , phenolic
Isocyanates eg. HNCO	NH <sub>2</sub>
Isothiocyanates R-NCS	NH <sub>2</sub>
Sulfonyl halides	NH <sub>2</sub>
Imidoesters	NH <sub>2</sub>
Diazoacetates	COOH, SH
Diazonium salts eg benzene-N2+ Cl- dicarbonyl compound	phenolic, imidazole NH-C(NH)-NH <sub>2</sub>

[0263] In one embodiment, the solid matrix may be housed in a flow chamber having an inlet and outlet to accommodate the multiple solutions and reactants that are allowed to flow past the immobilized capture probes. The flow chamber can be made of plastic or glass and may be either open or transparent in the plane viewed by a microscope or optical reader. Electro-osmotic flow includes a fixed charge on the solid support and a voltage gradient (current) passing between two electrodes placed at opposing ends of the solid support.

## Primers

[0264] In accordance with the invention, a primer is used to initiate replication by a polymerase of a target site on the target nucleic acid. If the polymerase is a DNA polymerase, the primer may be comprised of ribonucleotides or deoxyribonucleotides. The primers and/or the individual nucleotides or nucleotide analogs that are used to extend the primers may be suitably modified to enable signal generation, detection of the oligonucleotide products, and a determination of the presence or absence of the target sequence.

[0265] The primers used in the practice of the invention may be made synthetically, using conventional chemical or enzymatic nucleic acid synthesis technology. In one embodiment, the primers are less than about 25 nucleotides in length, from about 1 to about 10 nucleotides in length, or even about 2 to 3 nucleotides in length. It may be desirable to modify the nucleotides or phosphodiester linkages in one or more positions of the primer. Modifications may occur at 8-modified purines and 5-modified pyrimidines. Examples of such modifications include but are not limited to fluorescent molecules and energy transfer dyes (such as, fluorescein, aedans, coumarine, BODIPY dyes, and rhodamine based dyes), fluorescent quencher molecules (for example, Dabcyl), proteins, peptides, amino linkers, or amino acid based molecules (for example polyhistidine), modified bases and modified and unmodified base analogs, peptide nucleic acids (PNAs), methylphosphonates, radioactive labels, terminal phosphates, 3' glyceryl, other carbohydrate based molecules, fatty acid derived molecules, carbon spacer molecules, electrochemiluminescent labels, lanthanide labels, avidin and its derivatives (for example, streptavidin, Neutravidin, etc.), biotin, steroid molecules (such as Digoxygenin), thiol linkages, ferritin labels, and the like.

**Reaction Conditions**

[0266] Reaction conditions will vary depending on the template, polymerase, the reaction, and the nature of nucleotides and nucleotide analogs. Such conditions (e.g. salts, ionic strength, temperature, etc.) are known by those of ordinary skill in the art; and may be modified by routine experimentation to optimize condition, and to account for the presence of nucleotide analogs. Thiol-modified nucleotide analogues may be sensitive to strongly reducing conditions, such as high concentrations of 2-mercaptoethanol, and therefore lower concentrations of reducing agents may be appropriate and can be determined through routine experimentation. Reaction conditions may be designed for the production of full-length transcripts, or shorter length transcripts, such as abortive transcripts.

[0267] Examples of typical conditions and reagents for RNA polymerase transcription and DNA polymerase replication are readily found in the literature. See, e.g., Chamberlain et al., *The Enzymes*, Boyer, ed., New York Acad. Press, 3rd ed., p. 85 (1982); Dunn et al., *M. Mol. Biol.* 166: 477-535 (1983); Geider, *Proc. Natl. Acad. Sci. USA* 75: 645-649 (1978); Guruvich et al., *Analytical Biochem* 195: 207-213 (1991); Lewis et al., *J. Biol. Chem.* 255: 4928-4936 (1980); Martin et al., *Biochem.* 27: 3966-3974 (1988); and Milligan et al., *Methods Enzymol.* Vol. 180a, ed., 50-52 (1989)).

[0268] Chapter 11 of the well-known laboratory manual of Sambrook et al., *Molecular Cloning: A Laboratory Manual* (New York, Cold Spring Harbor Laboratory Press, 2001) describes hybridization conditions for oligonucleotide probes and primers in great detail, including a description of the factors involved and the level of stringency necessary to achieve hybridization with the desired degree of specificity.

**Application of nucleotide and nucleoside analogs.**

[0269] The uses and applications of nucleotide analogs, and nucleic acids containing such analogs are unlimited, and extend to all of those uses to which

detectably labelled compounds have been put in the art. For example, any compound to be detected and analyzed in a sample can be modified according to the techniques of the present invention. Labelled polynucleotides can be used in hybridization assays.

[0270] Nucleotide analogs are useful in the practice of any method utilizing nucleic acids. Such methods include, but are not limited to primer extension; terminal transferase additions; ligation; end labelling; polymerase chain reaction (PCR), ligation PCR, nick translation labelling, reverse transcription, Southern blotting, Northern blotting, ELISA, arrays, SKYE, cloning, transcription, abortive transcription, sequencing, diagnostic techniques, therapeutic techniques, and treatment and prevention of diseases and conditions.

[0271] In one embodiment of the invention are methods for the analysis of genetic material, such as DNA and RNA, wherein the genetic material is to be analyzed with a probe, such as a labeled probe having a nucleotide sequence complementary to the target genetic material. Such techniques comprise denaturing and fixing the target genetic material to a matrix. The target genetic material is then contacted with the denatured labeled probe so as to permit hybridization. After hybridization, those probes that did not hybridize to the target genetic material are separated from those probes that did hybridize to the target genetic material. Any method of separation can be utilized such as washing the target genetic material with a neutral solution to remove those probes that did not hybridize to the target genetic material. After the probe has been hybridized with the DNA or RNA to be identified, the detection system is added which can recognize the label of the probe and is capable of creating a detectable signal. The target genetic material may then be separated again, preferably by washing with a neutral solution, to remove those detection system moieties that did not recognize the label of the probe. Spectrophotometric or colorimetric techniques can be utilized to detect the hybridized probe.

[0272] The methods of the invention can be used in a variety of diagnostic contexts, such as methods of assessing the methylation state of specific genes, detecting the presence of known genetic mutations, detecting the presence of pathogenic organisms, detecting mRNA expression levels, and detecting and amplifying proteins.

[0273] Of interest is the detection and identification of viral and bacterial DNA sequences. Polynucleotide products prepared according to this invention can be utilized to diagnose genetic disorders by preparing a polynucleotide complementary to a DNA gene sequence which is associated with a genetic disorder, and detecting the presence of hybridization events. Detectable polynucleotides can also be used in chromosomal karyotyping, which comprises using a series of modified polynucleotides corresponding to a series of defined genetic sequences located on chromosomes, and then detecting hybridization.

[0274] Also of interest is the use of such compounds wherein a nucleic acid is coupled to antibodies, lectins, proteins, hormones and the like, and the presence of such nucleic acid forms a useful means for detecting the presence of such antibodies, lectins, proteins, hormones and the like. Thus, the present invention is useful in the modification of antibodies for use in immunoassay procedures, such as sandwich immunoassay procedures. Also of interest is the modification of drugs for assay procedures or of proteins associated with or known to be present on microorganism walls or membranes. Detectably labelled proteins prepared in such manner can also be used in competitive immunoassay procedures. Another use of the detectable compounds, especially biopolymers is in imaging, especially with monoclonal antibodies. These can be modified according to the techniques of the invention and allowed to carry a signal onto a given site in a living material, such as an animal body.

[0275] Nucleic acids containing modified analogs may be resistant to degradation or cutting with enzymes, such as restriction enzymes. In another embodiment, the nucleic acids may be more sensitive to degradation or

enzymatic activity. Differential stability between nucleic acids containing a modified nucleotide and those that do not contain such nucleotides may be used to distinguish newly synthesized nucleic acids from those present in a sample. In one embodiment, the random incorporation of nucleotide analogs in a nucleic acid may be used to create a large number of DNA fragments. Alternatively, greater stability of nucleic acids can be useful in methods, kits or therapy, as such species would be effective for longer periods of time than unmodified nucleic acid.

[0276] The invention also provides kits for carrying out the methods of the invention. Such kits comprise, in one or more containers, usually conveniently packaged to facilitate their use in assays, quantities of various compositions essential for carrying out the assays in accordance with the invention and appropriate instructions for use of the components in the containers.

[0277] Thus, the kits comprise one or more initiators according to the invention. The kits may additionally comprise an enzyme with polymerase activity, such as an RNA and/or DNA polymerase for example, to extend the primer of the kit, as well as reagents for processing a target nucleic acid. The kit may also comprise nucleotides and/or nucleotide analogs to enable detection of the oligonucleotide products synthesized by the methods of the invention. The kits may also include oligonucleotide probes. The kits may also contain components for the collection and transport of materials, including but not limited to, membranes, affinity materials, test tubes, petri dishes, and dipsticks. The kit may also include microtiter plates, bio-chips, magnetic beads, gel matrices, or other forms of solid matrices to which an oligonucleotide capture probe, which is specific for a particular target sequence, has been bound. The relative amounts of the components in the kits can be varied to provide for reagent concentrations that substantially optimize the reactions involved in the practice of the methods disclosed herein and/or to further optimize the sensitivity of any assay.

[0278] The test kits of the invention can also include, as is well-known to those skilled in the art, various controls and standards, such as solutions of known target nucleic acid concentration, including no target sequence (negative control), to ensure the reliability and accuracy of the assays carried out using the kits and to permit quantitative analyses of test samples using the kits.

[0279] The set of instructions are generally written instructions, though the instructions may be stored on electronic storage media (e.g., magnetic diskette or optical disk), relating to the use of the components of the methods of the invention. The instructions provided with the kit generally also include information regarding reagents (whether included or not in the kit) necessary or preferred for practicing the methods of the invention, instructions on how to use the kit, and/or appropriate reaction conditions.

[0280] One example is a kit for performance of abortive transcription on a target gene. Such a kit would include containers containing target site probe specific for the gene, an initiator, elongators, and a terminator, one of which is a nucleotide analog as described herein; RNA polymerase; a sample of the target gene as a positive control; and suitable instructions.

Abortive transcription.

[0281] In a number of embodiments of the invention, the nucleotide analogs and nucleic acids are used in methods associated with the technologies of abortive transcription (abscriptio) and reiterative oligonucleotide synthesis. These technologies are described in detail in U.S. Application No. 09/984,664, filed October 30, 2001; and International Application PCT/US02/34419, filed October 29, 2002; both of which are incorporated by reference herein.

Target Nucleic Acid for abortive transcription

[0282] The target nucleic acid is not a limiting factor in the methods of the invention, that have been further exemplified earlier. Thus, such nucleic acids

include both a naturally occurring or synthetic polynucleotide segment, and it can be obtained or synthesized by techniques that are well-known in the art.

#### Target Proteins

[0283] In another embodiment of the invention, the target may be another molecule, such as a protein, which is labeled by covalent or noncovalent attachment of a defined nucleic acid sequence which can be used for reiterative oligonucleotide synthesis (FIG. 21).

#### Immobilization

[0284] In one embodiment of the invention, the target molecule may be immobilized. In another embodiment, the target molecule may be immobilized to form, for example, a microarray. A single molecule array in accordance with this embodiment includes a solid matrix, a bioreactive or bioadhesive layer, and a bioresistant layer. Exemplifications suitable for the practice of the invention are described in greater detail herein.

#### Primers

[0285] In accordance with the invention, a primer is used to initiate replication by a DNA polymerase of a target site on the target nucleic acid. If the polymerase is a DNA polymerase, the primer may be comprised of ribonucleotides or deoxyribonucleotides. The primers and/or the individual nucleotides or nucleotide analogs that are used to extend the primers may be suitably modified to enable signal generation, detection of the oligonucleotide products, and a determination of the presence or absence of the target sequence.

[0286] The primers used in the practice of the invention may be made synthetically, using conventional chemical or enzymatic nucleic acid synthesis technology. In one embodiment, the primers are less than about 25 nucleotides in length, usually from about 1 to about 10 nucleotides in length,

and preferably about 2 to 3 nucleotides in length. In some embodiments of the invention the primer contains a modified nucleotide as described herein.

#### Target Site Probes

[0287] In accordance with the invention, an oligonucleotide target site probe is used to direct a polymerase to a target site on the target nucleic acid by forming a bubble complex in a target region of the target nucleic acid (FIG. 9). The target site probe may vary in the length of nucleotides, including but not limited to, about 20 to about 50 nucleotides, about 51 to about 75 nucleotides, about 76 to about 100 nucleotides, and greater than 100 nucleotides. The bubble complex comprises double-stranded regions on either side of a single-stranded region which includes a target site. In one embodiment, the target site probe includes three regions: a first region on the 5' end of the target site probe is complementary to and hybridizes with the template sequence upstream of a target site on the template sequence; a second region, which is 3' of the first region, is non-complementary to the template sequence and therefore does not hybridize with the template sequence; and a third region, which is on the 3' end of the target site probe, is complementary to and hybridizes with the template sequence downstream of the target site. The target site probe can vary in nucleotide length, including but not limited to, about 5-19; about 20 to about 50 nucleotides, about 51 to about 75 nucleotides, about 76 to about 100 nucleotides and greater than 100 nucleotides.

[0288] Use of the target site probe directs the polymerase to a particular enzyme binding site (*i.e.*, the double-stranded segment and bubble formed upstream of the target site by the template sequence and the primer) on the template sequence to facilitate the initiation of transcription at a particular target site. That is, rather than facilitating the random initiation of synthesis reactions by the polymerase along the length of a single-stranded template sequence, as described above, this embodiment provides targeted binding of

the polymerase for the detection of a particular target site encompassed by the bubble complex formed by the target site probe.

[0289] The target site probes used in the practice of the invention may be made enzymatically or synthetically, using conventional nucleic acid synthesis technology, such as phosphoramidite, H-phosphonate, or phosphotriester chemistry, for example. Alternative chemistries, such as those which result in non-natural backbone groups, such as phosphorothioate, phosphoramidate, and the like, may also be employed. The target site probes may be ordered commercially from a variety of companies which specialize in custom polynucleotides and/or oligonucleotides, such Operon, Inc. (Alameda, Calif.).

[0290] The sequence of the target site probe will vary depending upon the target sequence. The overall length of the target site probe is selected to provide for hybridization of the first and third regions with the target sequence and optimization of the length of the second, non-hybridized region. The first and third regions of the target site probe are designed to hybridize to known internal sites on the target nucleic acid template. Depending upon the application, the sequence of the second region on the target site probe can be designed such that the second region may or may not be self-complementary. The overall length of the target site probe ranges from about 20 to about 50 nucleotides, preferably from about 25 to about 35 nucleotides. The first and third regions of the target site probe each range from about 5 to about 20 nucleotides in length, preferably from about 8 to about 10 nucleotides in length. In one embodiment, the first and third regions of the target site probe are each about 10 nucleotides in length. The internal, second region on the target site probe ranges in length from about 8 to about 14 nucleotides, preferably from about 12 to about 14 nucleotides.

[0291] In one embodiment, at least one target site probe is used to specifically initiate abortive oligonucleotide synthesis at one or more target sites on the nucleic acid template to produce multiple oligonucleotide products. In another embodiment, the target site probe directs the initiation of abortive transcription on a single-stranded target site in the absence of a template promoter

sequence, as is known in the art. (See, U.S. Pat. No. 5,571,669; Daube and von Hippel, *Science*, 258: 1320-1324 (1992)).

#### Artificial Promoter Cassette

[0292] The term "Artificial Promoter Cassette" includes, in some embodiments, an "Abortive Promoter Cassette" as described in copending international application PCT/US02/34419, which is incorporated by reference herein. An artificial promoter cassette may be used for abortive or processive transcription depending on factors including the selection

[0293] of initiators, substrates, target size and reaction conditions.

[0294] In accordance with the invention, an artificial promoter cassette (APC) may be used to link a target to a defined sequence to generate multiple detectable oligonucleotide products that indicate the presence of the target in a test sample. The APC is a self-complementary sequence of DNA that may consist of: (1) one contiguous oligonucleotide to which RNA polymerase can bind to form a transcription bubble; (2) two partially complementary upper and lower oligonucleotides that form a single-stranded transcription bubble region comprising a defined site from which an initiator and a suitable RNA polymerase can synthesize an abortive oligonucleotide product; or (3) two complementary oligonucleotides that form a transcription bubble region in the presence of an RNA polymerase, which allows for the synthesis of an abortive oligonucleotide product. The APC may contain an artificial promoter, or it may contain the promoter for a specific RNA polymerase. For example, trinucleotide or tetranucleotide products that could be generated from with a common phage RNA polymerase can be made with a labeled GpA or GpApA initiator and a labeled pppG or pppA terminator.

[0295] In an exemplary embodiment, as illustrated in FIG. 1, the APC comprises eight regions, including an APC linker sequence which comprises either a 3' or a 5' single-stranded overhang region (*i.e.*, a "sticky end"). A first region (A) on the 5' end of the APC is complementary to a second region

(A') near the 3' end of the APC. A third region (B) and a fourth region (E) are separated from each other by regions C, D, and C' and are non-complementary to each other, such that the regions B and E form a single-stranded bubble region on the APC when the self-complementary regions of the APC interact with one another. Regions C and C' are substantially self-complementary, such that the 5' end of region C is complementary to the 3' end of the region C'. Region D may be a short sequence joining C and C' for a contiguous APC or may be a region comprising the free 3' or 5' ends of two separate upper and lower oligonucleotides for a two-part APC. Finally, the APC also includes an APC linker, a single-stranded region on either the 5' end or the 3' end of the APC oligonucleotide, which is formed through the complementary interaction of regions A and A'. The APC linker facilitates attachment of the APC with other target molecules, such as captured target DNA, RNA, or protein, for example.

[0296] The APC used in the practice of the invention may be made enzymatically or synthetically, using conventional nucleic acid synthesis technology, such as phosphoramidite, H-phosphonate, or phosphotriester chemistry, for example. Alternative chemistries, such as those that result in non-natural backbone groups, such as phosphorothioate, phosphoramidate, and the like, may also be employed. The APC may be ordered commercially from a variety of companies that specialize in custom polynucleotides and/or oligonucleotides, such as Operon, Inc. (Alameda, Calif.).

[0297] The length of the APC is selected to optimize the stability of the bubble region and provide for the hybridization of the APC linker sequence with the target sequence. The overall length of the APC may range from about 50 to about 150 nucleotides, preferably from about 55 to about 125 nucleotides. Regions A and A' may each comprise from about 5 to about 25 nucleotides and preferably comprise from about 7 to about 15 nucleotides. Regions B and E may comprise from about 8 to about 16 nucleotides and preferably comprise from about 10 to about 14 nucleotides. Regions C and C' may each comprise from about 5 to about 25 nucleotides and preferably

comprise from about 10 to about 20 nucleotides. The single-stranded overhang region may comprise from about 5 to about 40 nucleotides and preferably comprises from about 10 to about 25 nucleotides.

#### Polymerase

[0298] Template-dependent polymerases for use in the methods and compositions of the present invention are known in the art. The type of polymerase is not a limiting feature of the invention. Polymerases suitable for the practice of the invention are described in greater detail elsewhere within this application.

#### Nucleotides

[0299] In accordance with the invention, the polymerase catalyzes a reaction in the usual 5'→ 3' direction on the oligonucleotide product and either transcribes or replicates the target nucleic acid by extending the 3' end of the initiator or primer through the sequential addition of nucleotides (NTPs), which may include nucleotide analogs (NTP analogs) and which may be labeled or unlabeled. To facilitate reiterative, abortive synthesis initiation events, the NTPs and/or NTP analogs that are added to the reaction mixture before and/or during the synthesis reaction include a chain terminator, which is capable of terminating the synthesis event initiated by the polymerase. Use of the chain terminator stalls the polymerase during the synthesis reaction, inhibits formation of a processive elongation complex, and thereby promotes the reiterative synthesis of short abortive oligonucleotides from the target site. (Daube and von Hippel, *Science*, 258: 1320-1324 (1992)).

[0300] In accordance with the invention, a chain terminator may comprise any compound, composition, complex, reactant, reaction condition, or process (including withholding a compound, reactant, or reaction condition) which is capable of inhibiting the continuation of transcription or replication by the polymerase during the primer extension reaction. In one embodiment, a

suitable chain terminator is NTP deprivation, that is, depriving the polymerase of the particular NTP that corresponds to the subsequent complementary nucleotide of the template sequence. In other words, since NTP requirements for chain elongation are governed by the complementary strand sequence, given a defined template sequence and a defined primer length, a selected NTP may be withheld from the reaction mixture such that termination of chain elongation by the polymerase results when the reaction mixture fails to provide the polymerase with the NTP that is required to continue transcription or replication of the template sequence.

[0301] Alternatively, in another embodiment, the chain terminator may include nucleotide analogs, which may be labeled or unlabeled and which, upon incorporation into an oligonucleotide product by the polymerase, effect the termination of nucleotide polymerization. Specifically, since chain elongation by a polymerase requires a 3' OH for the addition of a subsequent nucleotide, nucleotide analogs having a suitably modified 3' end will terminate chain elongation upon incorporation into the oligonucleotide product. Nucleotide analogs having chain terminating modifications to the 3' carbon of the pentose sugar are known in the art and include nucleotide analogs such as 3' dideoxyribonucleoside triphosphates (ddNTPs) and 3' O-methylribonucleoside 5' triphosphates, as well as nucleotide analogs having either a —H or a —OCH<sub>2</sub> moiety on the 3' carbon of the pentose ring. Alternatively, in a further embodiment, the chain terminator may include nucleotide analogs, either labeled or unlabeled, which have a 3' OH group, but which, upon incorporation into the oligonucleotide product, still effect chain termination at some positions, as described herein (Costas, Hanna, et al., Nucleic Acids Research 28: 1849-58 (2000); Hanna, M., Meth Enzymology 180: 383-409 (1989); Hanna, M., Nucleic Acids Research 21: 2073-79 (1993); Hanna, M. et al., Nucleic Acid Research 27: 1369-76 (1999)).

[0302] NTPs and/or NTP analogs that can be employed to synthesize abortive oligonucleotide products in accordance with the methods of the invention may be provided in amounts ranging from about 1 to about 5000  $\mu$ M, preferably

from about 10 to about 2000  $\mu$ M. In a preferred aspect, nucleotides and/or nucleotide analogs, such as ribonucleoside triphosphates or analogs thereof, that can be employed to synthesize oligonucleotide RNA transcripts by the methods of the invention may be provided in amounts ranging from about 1 to about 6000  $\mu$ M, preferably from about 10 to about 5000  $\mu$ M.

#### Labeling and Detection

[0303] In accordance with an aspect of the invention, detectable oligonucleotide products are synthesized from a target nucleic acid template. The detection and identification of the oligonucleotide products are facilitated by label moieties on the initiator and/or on the NTPs or NTP analogs that are incorporated by the polymerase into each oligonucleotide product that is synthesized on the target nucleic acid and/or on other molecules which are part of the synthetic complex or which interact with one or more components of the synthetic complex. The label or reporter moieties may be chemically or enzymatically incorporated into the nucleotides forming the primer and/or into the reactant NTPs or NTP analogs that are utilized by the polymerase during the extension reaction, or other molecules. Examples of appropriate labelled nucleotide analogs are described elsewhere within this application.

[0304] In accordance with the invention, detection of the oligonucleotide products is indicative of the presence of the target sequence. Quantitative analysis is also feasible. Direct and indirect detection methods (including quantitation) are well known in the art. For example, by comparing the amount of oligonucleotide products that are generated from a test sample containing an unknown amount of a target nucleic acid to an amount of oligonucleotide products that were generated from a reference sample that has a known quantity of a target nucleic acid, the amount of a target nucleic acid in the test sample can be determined. The reiterative abortive synthesis initiation and detection methods of the present invention can also be extended

to the analysis of genetic sequence alterations in the target nucleic acid, as further described below.

#### Reaction Conditions

[0305] Most transcription reaction conditions are designed for the production of full length transcripts, although no conditions have been identified that eliminate abortive transcription. Appropriate reaction media and conditions for carrying out the methods of the present invention include an aqueous buffer medium that is optimized for the particular polymerase. In general, the buffer includes a source of monovalent ions, a source of divalent cations, and a reducing agent, which is added to maintain sulfhydryl groups in the polymerase in a reduced form. Any convenient source of monovalent ions, such as KCl, K-acetate, NH<sub>4</sub>-acetate, K-glutamate, NH<sub>4</sub>Cl, ammonium sulfate, and the like, may be employed. The divalent cation may be magnesium, manganese, zinc, or the like, though, typically, the cation is magnesium (Mg). Any convenient source of magnesium cations may be employed, including MgCl<sub>2</sub>, Mg-acetate, and the like. The amount of Mg<sup>2+</sup> present in the buffer may range from about 0.5 to 20 mM, preferably from about 1 to 12 mM.

[0306] Representative buffering agents or salts that may be present in the buffer include Tris Phosphate, Tricine, HEPES, MOPS, and the like, where the amount of buffering agent typically ranges from about 5 to 150 mM, usually from about 10 to 100 mM, and preferably from about 20 to 50 mM. In certain embodiments, the buffering agent is present in an amount sufficient to provide a pH ranging from about 6.0 to 9.5, preferably ranging from about 7.0 to 8.0. Other agents which may be present in the buffer medium include chelating agents, such as EDTA, EGTA, and the like, or other polyanionic or cationic molecules (heparins, spermidine), protein carriers (BSA) or other proteins, including transcription factors (sigma, NusA, Rho, lysozyme, GreA, GreB, NusG, etc.).

[0307] Variations in all of the reaction components potentially can alter the ratio of abortive transcripts to full-length transcripts. Alterations in the concentration of salts (from 10 mM to 100 mM) or the use of alternative monovalent cations ( $K^+$  versus  $Na^+$  versus  $Rb^+$ ) have been shown to affect the level of transcription (measured as abortive transcription) on linear DNA templates (Wang, J-Y, *et al.*, *Gene* 196:95-98 (1997)). Alternative sulfhydral reducing reagents are reported to have differential effects on abortive transcription. 2-mercaptoethanol at 1-2mM is reported to enhance abortive transcription on a poly[dA-dT] template compared to the alternative reducing agent 5,5'-dithio-bis-(2-nitrobenzoic) acid (Job, D., *Acta Biochem. Pol.* 41:415-419 ((1994))).

[0308] A high molar ratio of RNA polymerase to template enhances the frequency of abortive transcription over full length transcription on the lambda  $P_R$  promoter. This effect apparently arises from collisions between tandem polymerases at the promoter.

[0309] Certain RNA polymerase mutants have elevated rates of abortive transcription compared to the wild-type polymerase. For example, a mutation changing an arginine to a cysteine at codon 529 in the RNA polymerase beta subunit gene causes elevated abortive transcription at the *E. coli* pyrB1 promoter (Jin, D.J. and Turnbough, Jr., C.L., *J. Mol. Biol.* 236:72-80 (1994)).

[0310] The relative level of abortive transcription is sensitive to the nucleotide sequence of the promoter. A number of promoters have been identified that are unusually susceptible to abortive transcription (*e.g.*, the galP2 promoter). The assay system that relies on recruitment of a defined promoter can be optimized by screening candidate promoters for maximal initiation frequency and maximal proportion of abortive transcripts.

[0311] Any aspect of the methods of the present invention can occur at the same or varying temperatures. In one embodiment, the reactions are performed isothermally, which avoids the cumbersome thermocycling process. The synthesis reaction is carried out at a temperature that permits hybridization of the various oligonucleotides, including target site probes,

capture probes, and APCs, as well as the primers to the target nucleic acid template and that does not substantially inhibit the activity of the enzymes employed. The temperature can be in the range of about 25 °C to about 85 °C, more preferably about 30 °C to about 75 °C, and most preferably about 25 °C to about 55 °C. In some embodiments, the temperature for the transcription or replication may differ from that temperature(s) used elsewhere in the assay. The temperature at which transcription or replication can be performed is in the range of about 25 °C to about 85 °C, more preferably about 30 °C to about 75 °C, and most preferably about 25 °C to about 55 °C.

[0312] Denaturation of the target nucleic acid in a test sample may be necessary to carry out the assays of the present invention in cases where the target nucleic acid is found in a double-stranded form or has a propensity to maintain a rigid structure. Denaturation is a process that produces a single-stranded nucleic acid and can be accomplished by several methods that are well-known in the art. See, e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual (New York, Cold Spring Harbor Laboratory Press, third edition, 2000). One method for achieving denaturation includes the use of heat, such as exposing the nucleic acid in a test sample to temperatures of about 90-100 °C for about 2-20 minutes. Alternatively, a base may be used as a denaturant when the nucleic acid comprises DNA. Many basic solutions, which are well-known in the art, may be used to denature a DNA sample. An exemplary method incubates the DNA sample with a base, such as NaOH for example, at a concentration of about 0.1 to 2.0 N NaOH at a temperature ranging from about 20 °C to about 100 °C for about 5-120 minutes. Treatment with a base, such as sodium hydroxide, not only reduces the viscosity of the sample, which increases the kinetics of subsequent enzymatic reactions, but also aids in homogenizing the sample and reducing the possibility of background by destroying any existing DNA-RNA or RNA-RNA hybrids that may exist in the sample.

[0313] In accordance with various aspects and embodiments of the invention, the target nucleic acid molecules may be hybridized to an oligonucleotide

capture probe, a mononucleotide or oligonucleotide initiator which is complementary to a portion of the target nucleic acid, an APC linker sequence that is complementary to a portion of a target nucleic acid, and/or a target site probe that is complementary to regions on either side of the target site. Hybridization is conducted under standard hybridization conditions that are well-known to those skilled in the art. Reaction conditions for hybridization of an oligonucleotide (or polynucleotide) to a target sequence vary from oligonucleotide to oligonucleotide, depending upon factors such as oligonucleotide length, the number of G:C base pairs present in the oligonucleotide, and the composition of the buffer utilized in the hybridization reaction. Moderately stringent hybridization conditions are generally understood by those skilled in the art to be conditions that are approximately 25 °C below the melting temperature of a perfectly base-paired double-stranded DNA. Higher specificity is generally achieved by employing more stringent conditions, such as incubation conditions having higher temperatures. Chapter 11 of the well-known laboratory manual of Sambrook et al., *Molecular Cloning: A Laboratory Manual* (New York, Cold Spring Harbor Laboratory Press, 1989) describes hybridization conditions for oligonucleotide probes and primers in great detail, including a description of the factors involved and the level of stringency necessary to achieve hybridization with the desired degree of specificity.

[0314] The oligonucleotide capture probe, the target site probe, the APC, and/or the initiator may each be incubated with the target nucleic acid for about 5 to 120 minutes at about 20 to 80 °C to permit hybridization. Preferably, the target nucleic acid and the oligonucleotide probes, the APC, and/or the initiator are incubated for about 5 to 60 minutes at about 25 to 70 °C. More preferably, the target nucleic acid and the oligonucleotide probes, the APC, and/or primer are incubated for about 5-30 minutes at about 35-50 °C.

[0315] Hybridization is typically performed in a buffered aqueous solution and temperature conditions, salt concentration, and pH are selected to provide

sufficient stringency to enable the oligonucleotide probes, the APC, or the primer to hybridize specifically to the target sequence but not to any other sequence. Generally, the efficiency of hybridization between an oligonucleotide or polynucleotide and a target nucleic acid template will be improved under conditions where the amount of oligonucleotide or polynucleotide added to the reaction mixture is in molar excess to the template, preferably a molar excess that ranges from about 10<sup>3</sup> to 10<sup>6</sup>. It will be appreciated, however, that the amount of target nucleic acid in the test sample may not be known, so that the amount of an oligonucleotide, such as the amount of an oligonucleotide capture probe, a target site probe, or an APC for example, relative to an amount of a target nucleic acid template cannot be determined with certainty.

[0316] Alternatively, if a target DNA sequence has been treated with a base to effect denaturation, the oligonucleotide or polynucleotide is diluted in a probe diluent that also acts as a neutralizing hybridization buffer. In this manner, the pH of the test sample can be kept between about 6 and 9, which will favor the hybridization reaction and will not interfere with subsequent enzymatic reactions. Preferably, the neutralizing buffer is a 2-[bis(2-hydroxyethyl) amino] ethane sulfonic acid ("BES") (Sigma, St. Louis, Mo.) and sodium acetate buffer. More preferably, the neutralizing hybridization buffer is a mixture of 2 M BES, 1 M sodium acetate, 0.05% of an antimicrobial agent, such as NaN<sub>3</sub>, 5 mM of a chelating agent, such as EDTA, 0.4% of a detergent, such as Tween-20<sup>TM</sup>, and 20% of a hybridization accelerator, such as dextran sulfate. The pH of the neutralizing hybridization buffer is between about 5 to 5.5.

[0317] Transcription conditions and reagents are well-known in the art. Examples of typical conditions and reagents for RNA polymerase transcription and DNA polymerase replication are readily found in the literature. See, e.g., Chamberlain et al., *The Enzymes*, Boyer, ed., New York Acad. Press, 3rd ed., p. 85 (1982); Dunn et al., *M. Mol. Biol.* 166: 477-535 (1983)); Geider, *Proc. Natl. Acad. Sci. USA* 75: 645-649 (1978)); Guruvich et

al., Analytical Biochem 195: 207-213 (1991); Lewis et al., J. Biol. Chem. 255: 4928-4936 (1980); Martin et al., Biochem. 27: 3966-3974 (1988); and Milligan et al., Methods Enzymol. Vol. 180a, ed., 50-52 (1989)). As described in Lu et al., U.S. Patent No. 5,571,669, polymerase concentrations for transcription initiated from artificial transcription bubble complexes are generally about one order of magnitude higher than the ideal polymerase concentrations for promoter-initiated, or palindromic sequence-initiated, transcription.

[0318] In one embodiment, the foregoing components are added simultaneously at the initiation of the abortive synthesis and detection methods. In another embodiment, components are added in any order prior to or after appropriate timepoints during the method, as required and/or permitted by the various reactions. Such timepoints can be readily identified by a person of skill in the art. The enzymes used for nucleic acid detection according to the methods of the present invention can be added to the reaction mixture prior to or following nucleic acid denaturation, prior to or following hybridization of the primer to the target nucleic acid, prior to or following the optional hybridization of the target site probe to the target nucleic acid, or prior to or following the optional hybridization of the APC, as determined by the enzymes' thermal stability and/or other considerations known to those skilled in the art.

[0319] The various reactions in the methods of the invention can be stopped at various timepoints and then resumed at a later time. These timepoints can be readily identified by a person of skill in the art. Methods for stopping the reactions are known in the art, including, for example, cooling the reaction mixture to a temperature that inhibits enzyme activity. Methods for resuming the reactions are also known in the art, including, for example, raising the temperature of the reaction mixture to a temperature that permits enzyme activity. In some embodiments, one or more of the components of the various reactions may be replenished prior to, at the time of, or following the

resumption of the reactions. Alternatively, the reaction can be allowed to proceed (*i.e.*, from start to finish) without interruption.

#### Abortive Synthesis and Detection Methods of the Invention

[0320] The following examples of the abortive synthesis and detection methods of the invention are provided to more specifically describe the invention. These exemplary methods are intended to be merely illustrative and are not intended to limit the description provided above. It will be appreciated that various other embodiments may be practiced, given the above general description. For example, reference to the use of a primer means that any of the primers described herein may be used, including RNA initiators.

[0321] In accordance with an aspect of the invention, a method for detecting the presence of a target polynucleotide by generating multiple detectable oligonucleotide products through reiterative synthesis initiation events on the target polynucleotide is provided. FIG. 2 diagrammatically illustrates the various reactants that may be combined and reacted in the presence of RNA polymerase to synthesize multiple detectable oligonucleotide products. The methods of the invention may be performed using a test sample that potentially contains a target sequence. The test sequence may be detected directly or the product of primer-extension or reverse transcription of the target may be detected. Sequences or tags may be added to the copy of the target (*e.g.*, biotin, ssDNA regions). The test sample may include double-stranded DNA, single-stranded DNA, or RNA. The DNA or RNA may be isolated and purified by standard techniques for isolating DNA or RNA from cellular, tissue, or other samples. Such standard methods may be found in references such as Sambrook et al., *Molecular Cloning: A Laboratory Manual* (New York, Cold Spring Harbor Laboratory Press, third edition, 2000). In one embodiment, the target nucleic acid is DNA or RNA that is in a suitable medium, although the target nucleic acid can be in lyophilized form. Suitable media include, but are not limited to, aqueous media (such as pure water or

buffers). In another embodiment, the target nucleic acid is immobilized prior to being utilized as a substrate for a synthesis reaction.

[0322] In an exemplary embodiment, the target sequence is immobilized by a sequence-specific (e.g., gene-specific) oligonucleotide capture probe that is attached to a solid matrix, such as a microtiter plate. The immobilized capture probe is treated under hybridizing conditions with a test sample that includes single-stranded DNA (i.e., denatured DNA) or RNA. Any target sequence that is present in the test sample hybridizes to the capture probe and is then exposed to additional reagents in accordance with the invention.

[0323] In an exemplary embodiment, an initiator (n 5'-R<sub>1</sub>-(N<sub>1</sub>)<sub>x</sub>-OH 3') hybridizes with the target sequence upstream of a target site in the presence of the target site probe (FIG. 9) and facilitates catalysis of a polymerization reaction at the target site by the polymerase. The initiation primer may be comprised of nucleosides, nucleoside analogs, nucleotides, and nucleotide analogs. The initiator primer may vary in the number of nucleotides, such as nucleotides from 1-25 nucleotides, 26-50 nucleotides, 51-75 nucleotides, 76-100 nucleotides, 101-125 nucleotides, 126-150 nucleotides, 151-175 nucleotides, 176-200 nucleotides, 201-225 nucleotides, 226-250 nucleotides, and greater than 250 nucleotides, and may include one or more nucleotide analogs. A suitable RNA polymerase is employed to synthesize an oligoribonucleotide product from the target sequence or any portion thereof. The polymerase may be an RNA-dependent or DNA-dependent RNA polymerase. The DNA or RNA target sequence may or may not be attached to other molecules, such as proteins, for example.

[0324] During the polymerization reaction, the initiator is extended or elongated by the polymerase through the incorporation of nucleotides which have been added to the reaction mixture. As the polymerase reaction proceeds, the polymerase extends the initiator, as directed by the template sequence, by incorporating corresponding nucleotides that are present in the reaction mixture. In one embodiment, these reactant nucleotides comprise a chain terminator (e.g., n 5' pppN<sub>T</sub>-R<sub>2</sub>, a chain terminating nucleotide analog,

as described above). When the polymerase incorporates a chain terminator into the nascent oligonucleotide product, chain elongation terminates due to the polymerase's inability to catalyze the addition of a nucleotide at the 3' position on the pentose ring of the chain terminator. Consequently, the polymerase aborts the initiated synthesis event by releasing the oligonucleotide product (*i.e.*, 5' R<sub>1</sub>-(N<sub>1</sub>)<sub>z</sub>pN<sub>T</sub>-R<sub>2</sub>, where z = x + y) and reinitiating the abortive initiation synthesis reaction at the target site.

[0325] The abortive initiation reaction may be controlled such that the polymerase aborts synthesis after extending the initiator by a predetermined number of nucleotides. For example, if it is desirable to terminate the synthesis reaction after the initiator has been extended by a single nucleotide, this may be accomplished by, for example, either: (1) adding to the reaction mixture only nucleotides that are chain terminators, thereby inhibiting polymerization after the first nucleotide is incorporated by the polymerase; or (2) if the genetic sequence of the target site is known, adding to the reaction mixture only a preselected chain terminating nucleotide analog (*i.e.*, nucleotide analogs which comprise one of A, G, T, C, or U) that is complementary to the nucleotide at the target site. Alternatively, if it is desirable to terminate the synthesis reaction after the initiator has been elongated by a predetermined number of nucleotides, and if the genetic sequence of the target site is known, this may be accomplished by, for example, adding to the reaction mixture a preselected chain terminating nucleotide analog (*i.e.*, nucleotide analogs which comprise one of A, G, T, C, or U) that is complementary to an Nth nucleotide from the target site, where N is the predetermined number of nucleotides comprised by the oligonucleotide product, exclusive of the initiator. In this manner, multiple abortive oligonucleotide products that comprise the initiator and a chain terminating nucleotide analog are synthesized by the polymerase.

[0326] The polymerase releases the oligonucleotide product without translocating from the enzyme binding site or dissociating from the target polynucleotide sequence. Nucleotide deprivation can be used to sequester the

polymerase at the polymerase binding site. For example, if only an initiator and a terminator are supplied, elongation by the polymerase will not be possible.

[0327] Furthermore, reaction conditions may be optimized for abortive transcription initiation, whereby it is favorable for the polymerase to remain bound to the polymerase binding site even in the presence of elongating nucleotides. The abortive initiation reaction buffer will be optimized to increase the abortive events by adjusting the concentrations of the salts, the divalent cations, the glycerol content, and the amount and type of reducing agent to be used. In addition, "roadblock" proteins may be used to prevent the polymerase from translocating.

[0328] In another aspect of the invention, the initiator includes a moiety (e.g., R<sub>1</sub>, as depicted in FIG. 2) which may be covalently bonded to the 5' phosphate group, the 2' position of the pentose ring, or the purine or pyrimidine base of one of the nucleotides or nucleotide analogs that are incorporated into the initiator. Additionally, the reactant nucleotides and/or nucleotide analogs that are included in the reaction mixture for incorporation into the oligonucleotide product by the polymerase may each also include a moiety (e.g., R<sub>2</sub>, as depicted in FIG. 2), which is covalently bonded to either the nucleobase (as in FIG. 3) or the 2' position or 3' position of the pentose ring. In an exemplary embodiment, R<sub>1</sub> and R<sub>2</sub> are label moieties (as in FIG. 4) on the initiator and the chain terminator, respectively, that are incorporated into the oligonucleotide product by the polymerase (as in FIG. 5) and are adapted to interact in a manner that generates a detectable signal (e.g., fluorescence resonance energy transfer (FRET), fluorescence or colorimetry (FIG. 6)), thereby permitting the detection and quantitation of the synthesized oligonucleotide products. In one embodiment, as illustrated in FIG. 7, an oligonucleotide product (5' R<sub>1</sub>- (N<sub>I</sub>)<sub>x</sub>pN<sub>T</sub> - R<sub>2</sub>) incorporating an initiator (N<sub>I</sub>) that has an energy donor group (R<sub>1</sub>) and a chain terminating nucleotide (N<sub>T</sub>) that has an energy acceptor group (R<sub>2</sub>) generates a signal through fluorescence resonance energy transfer from R<sub>1</sub> to R<sub>2</sub> when the synthesized oligonucleotide

products are irradiated with light of a particular wavelength. As shown in FIG. 9, when the energy donor moiety  $R_1$  on the initiator is excited by exposure to light of a specified wavelength ( $\lambda_{1A}$ ) (e.g., the absorption maximum of  $R_1$ ) the excited donor moiety  $R_1$  emits light of a second wavelength ( $\lambda_{1E/2A}$ ) (e.g., the emission maximum for  $R_1$ ) which is absorbable by  $R_2$ . If  $N_T$  has been suitably incorporated into the oligonucleotide product by the polymerase, the energy acceptor moiety  $R_2$  on  $N_T$  is positioned sufficiently near  $R_1$  on  $N_I$  (e.g., within about 80 Å) to facilitate efficient energy transfer between  $R_1$  and  $R_2$ , such that  $R_2$  absorbs the wavelength of light ( $\lambda_{1E/2A}$ ) emitted by the excited donor moiety  $R_1$ . In response to the absorption of  $\lambda_{1E/2A}$ , the excited  $R_2$  acceptor moiety emits light of a third wavelength ( $\lambda_{2E}$ ), which may then be detected and quantified in accordance with methods that are well-known in the art. Exemplary  $R_1$  and/or  $R_2$  FRET label moieties include aedans and fluorescein, or pyrene, stilbene, coumarine, bimane, naphthalene, pyridyloxazole, naphthalimide, NBD, BODIPY™, as well as any of those described in greater detail above.

[0329] In an alternate embodiment, as diagrammatically illustrated in FIG. 8,  $n$  copies of a dinucleotide initiator (5'  $R_1-N_1pN_2-R_2-OH$  3') comprising reporter moieties ( $R_1$  and  $R_2$ ) on either of the nucleotides ( $N_1$  and  $N_2$ , respectively) may be extended by a polymerase to incorporate  $n$  copies of a chain terminator (5' ppp $N_3R_3$ ) which includes a third reporter moiety ( $R_3$ ), yielding  $n$  copies of a detectable trinucleotide transcript (5'  $R_1-N_1pN_2R_2pN_3-R_3-OH$  3'). In a manner similar to the one described above with reference to FIG. 8, the trinucleotide transcript may be irradiated with a first wavelength of light ( $\lambda_{1A}$ ) which excites the  $R_1$  energy donor group on the first nucleotide ( $N_1$ ) to emit  $\lambda_{1E}/\lambda_{3A}$ .  $\lambda_{1E}/\lambda_{3A}$  is then absorbed by the  $R_3$  energy acceptor group on the chain terminating nucleotide ( $N_3$ ), and an excited  $R_3$  then emits  $\lambda_{3E}$ , which can then be detected and quantified. Alternatively, the transcript may be irradiated with a second wavelength of light ( $\lambda_{2A}$ ) which excites an  $R_2$  energy donor group on a second nucleotide ( $N_2$ ) to emit  $\lambda_{2E}/\lambda_{3A}$ .  $\lambda_{2E}/\lambda_{3A}$  is

then absorbed by the  $R_3$  energy acceptor group, and an excited  $R_3$  then emits  $\lambda_{3E}$ , which can be detected and quantified. In either case, the detectable wavelength ( $\lambda_{3E}$ ) is not obtained unless the polymerase brings an energy donor reporter moiety on the initiator ( $R_1$  or  $R_2$ ) into sufficient proximity with a corresponding energy acceptor reporter moiety ( $R_3$ ) on the incorporated nucleotide to result in the emission of the detectable wavelength of light.

[0330] In another aspect of the invention, as diagrammatically illustrated in FIG. 9, a target site probe may be used to form a bubble complex in a target region of the target sequence. As described above, the bubble complex comprises double-stranded regions that flank a single-stranded region that includes a target site. In this embodiment, the target site probe is used to direct the polymerase to the target site by positioning the target site at the junction of the single-stranded bubble region and a downstream duplex region on the target sequence. In an exemplary embodiment, the target site probe comprises from about 18-54 nucleotides: a first region (A) which hybridizes to the target sequence (A') upstream of the target site comprises about 5-20 nucleotides; an internal, second region of non-base-paired nucleotides (B) comprises about 8-14 nucleotides; and a third region (C) which hybridizes to the target sequence downstream of the target site (C') comprises about 5-20 nucleotides. The polymerase associates with an initiator and initiates a synthesis reaction at the target site on the template sequence. The polymerase elongates the initiator to synthesize an abortive oligonucleotide product through the incorporation of nucleotides, which comprise a suitable chain terminator. Both the initiator and the nucleotides, including the chain terminating nucleotide, may be modified with a label moiety to allow signal detection, such as by fluorescence resonance energy transfer for example, as described above.

[0331] An illustrative procedure for detecting multiple oligonucleotide products through reiterative synthesis initiation events on a target sequence, therefore, may include: (a) optionally immobilizing an oligonucleotide capture probe which is designed to hybridize with a specific or general target

sequence; (b) optionally hybridizing the oligonucleotide capture probe with a test sample which potentially contains a target sequence; (c) optionally hybridizing the target sequence with a target site probe; (d) modifying at least one of an initiator and nucleotides comprising a chain terminator to enable detection of the oligonucleotide product synthesized by the polymerase; (e) hybridizing the target sequence with the primer; and (f) extending the initiator with a polymerase such that the polymerase reiteratively synthesizes an oligonucleotide product that is complementary to a target site by incorporating complementary nucleotides comprising a chain terminator and releasing an abortive oligonucleotide product without either translocating from an enzyme binding site or dissociating from the target sequence.

[0332] During transcription of the template by the RNA polymerase, the RNA initiator is extended by the RNA polymerase through the incorporation of nucleotides that have been added to the reaction mixture. As the polymerase reaction proceeds, the RNA polymerase extends the RNA initiator, as directed by the template sequence, by incorporating corresponding nucleotides that are present in the reaction mixture. In one embodiment, these reactant nucleotides comprise a chain terminator (e.g., a 5'  $\text{pppN}_T\text{-R}_2$ , a chain terminating nucleotide analog, as described above). When the RNA polymerase incorporates a chain terminator into the nascent transcript, chain elongation terminates due to the polymerase's inability to catalyze the addition of a nucleotide at the 3' position on the ribose ring of the chain terminator, and the RNA polymerase aborts the initiated transcription event by releasing the transcript and reinitiating transcription at the target site. The abortive transcription initiation reaction may be controlled such that multiple abortive oligonucleotide transcripts of a predetermined length and comprising the RNA primer and a chain terminating nucleotide analog are generated.

[0333] In an exemplary embodiment, the RNA initiator may be a mononucleotide and the nucleotides provided in the reaction mixture may comprise solely chain terminators. In this embodiment, transcription is aborted by the RNA polymerase after the RNA initiator has been extended by

a single nucleotide and an abortive dinucleotide transcript is generated. In another embodiment, the RNA initiator may comprise a dinucleotide or a trinucleotide, for example, and an abortive transcription initiation event may generate an abortive transcript comprising a trinucleotide or a tetranucleotide, respectively. It will be appreciated that abortive transcripts of any desired length may be obtained, depending upon the length of the RNA initiator and the nature and composition of the reactant nucleotides that are selected for inclusion in the reaction mixture. For example, if the nucleotide sequence of the template is known, the components (e.g., target site, initiator, and reactant nucleotides) of the transcription reaction may be selected such that abortive transcripts of any desired length are generated by the method of the invention.

[0334] In another aspect of the invention, the RNA initiator includes a moiety (e.g., R<sub>1</sub>, as depicted in FIG. 5) which may be covalently bonded to the 5' phosphate group, the 2' position of the ribose ring, or the purine or pyrimidine base of one of the nucleotides or nucleotide analogs that are incorporated into the RNA initiator. Additionally, the reactant nucleotides and/or nucleotide analogs that are included in the reaction mixture for incorporation into the oligonucleotide transcript by the RNA polymerase may each also include a moiety (e.g., R<sub>2</sub>, as depicted in FIG. 5), which is covalently bonded to either the nucleobase or the 2' position or 3' position of the ribose ring. The moieties R<sub>1</sub> and R<sub>2</sub> may each comprise H, OH, or any suitable label moiety, reporter group, or reporter group precursor, as described in greater detail above.

[0335] An illustrative procedure for detecting multiple oligonucleotide transcripts through reiterative transcription initiation events on a target sequence, therefore, may include: (a) optionally immobilizing an oligonucleotide capture probe which is designed to hybridize with a specific or general target sequence; (b) optionally hybridizing the oligonucleotide capture probe with a test sample which potentially contains a target sequence; (c) optionally hybridizing the target sequence with a target site probe; (d) modifying at least one of an RNA initiator and nucleotides comprising a chain

terminator to enable detection of the oligonucleotide transcript synthesized by the RNA polymerase; (e) hybridizing the target sequence with the RNA initiator; and (f) extending the RNA initiator with an RNA polymerase such that the RNA polymerase reiteratively synthesizes an oligonucleotide transcript that is complementary to a target site by incorporating complementary nucleotides comprising a chain terminator and releasing an abortive oligonucleotide transcript without substantially translocating from the polymerase binding site or dissociating from the target sequence.

[0336] In accordance with another aspect of the invention, as diagrammatically illustrated in FIG. 6, the methods of the invention may be utilized to generate an oligonucleotide product (5' R<sub>1</sub>-(N<sub>1</sub>)<sub>x</sub>pN<sub>T</sub>-R<sub>2</sub>) which comprises an initiator (N<sub>1</sub>) with a moiety (R<sub>1</sub>), such as an immobilization tag for example; and a chain terminating nucleotide (N<sub>T</sub>) that includes a label moiety (R<sub>2</sub>), such as a signal generator or signal generator precursor for example. In this embodiment, the oligonucleotide product(s) may be captured or immobilized, such as on a membrane for example, to facilitate detection of the oligonucleotide products of the abortive synthesis reaction. In an exemplary embodiment, R<sub>1</sub> is a bioadhesive tag, such as biotin for example; R<sub>2</sub> is a label moiety, such as fluorescein for example; and oligonucleotide products that are attached to the solid matrix by the R<sub>1</sub> bioadhesive tag are capable of direct detection through an emission from the R<sub>2</sub> label moiety. In another exemplary embodiment, an antibody, such as anti-dinitrophenyl (anti-DNP) for example, is attached to the solid matrix; R<sub>1</sub> is an immobilization tag, such as dinitrophenyl (DNP) for example; R<sub>2</sub> is a reporter or reporter precursor, such as a reactive thiol for example; and, upon silver/gold development, the oligonucleotide products that are attached to the solid matrix by the R<sub>1</sub> tag produce a colored signal that is visible to the naked eye without irradiation.

Applications of the Abortive Synthesis and Detection Methods of the Invention

[0337] The methods of the present invention can be used in a variety of diagnostic contexts. For purposes of illustration, methods of assessing the methylation state of specific genes, detecting the presence of known genetic mutations, detecting the presence of pathogenic organisms, detecting mRNA expression levels, and detecting and amplifying proteins are described.

DNA Methylation

[0338] The methods of the present invention may be used in diagnostic assays which detect epigenetic changes associated with disease initiation and progression by assessing the methylation state of specific genes and their regulatory regions that are known to be associated with particular disease-states. DNA methylation is a cellular mechanism for altering the properties of DNA without altering the coding function of that sequence. The methylation reaction, which is catalyzed by DNA-(cystosine-5)-methyltransferase, involves the transfer of a methyl group from S-adenosylmethionine to the target cytosine residue to form 5-methylcytosine (5-mCyt) (FIG. 10). See Gonzalgo et al., U.S. Pat. No. 6,251,594. The areas of the genome that contain 5-mCyt at CpG dinucleotides are referred to as "CpG islands." While changes in the methylation status of the cytosine residues in DNA CpG islands commonly occur in aging cells, altered gene methylation (either increased or decreased) is frequently an early and permanent event in many types of disease, including cancer. CpG islands tend to be found in DNA regulatory regions that are near genes and determine whether these genes are either active or inactive. Many genes that regulate cell growth, and therefore prevent or inhibit the development of cancer, such as tumor suppressor genes, must be active (unmethylated) to promote normal cell growth. Other genes, such as oncogenes for example, must be inactive (methylated) so as not to promote abnormal cell growth.

[0339] For example, many types of cancer are associated with a distinct combination or pattern of CpG island methylation. FIG. 14 graphically illustrates the manner in which altered gene methylation may be associated with various types of cancer. The graph plots 13 exemplary cancers (prostate, kidney, bladder, esophageal, lung, gastric, colon, blood, breast, skin, brain, liver, and ovarian) against 49 genes which have been shown to have methylation changes that are associated with the initiation and progression of the identified types of cancer. Each oval in the graph (coded by cancer type) indicates an abnormal methylation status for a gene (*i.e.*, methylated when its normal status is unmethylated or unmethylated when its normal status is methylated). Since each type of cancer may be associated with a different pattern of methylation-altered genes, cancer-affected organs may potentially be identified based upon organ-specific combinations of methylated genes. For example, in the case of prostate cancer cells, genes 4, 9, 10, 14, 19, 22, 32, and 33 have been shown to exhibit abnormal methylation states. Thus, if standardized diagnostics could easily evaluate the methylation states of these 8 genes, then the initiation, progression, and recurrence of prostate cancer could be readily monitored and more effective patient treatment strategies could be developed. It will be appreciated that FIG. 14 represents only a subset of the genes for which altered methylation states and patterns are indicative of various types of cancer.

#### Cancer as a genetic disease

[0340] Cancer is actively prevented through the expression of numerous tumor-suppressor genes that regulate the cell-division cycle. This negative control system balances the effects of growth promoting genes (proto-oncogenes) to ensure that individual cells undergo growth and division only under highly defined conditions. Studies of benign and malignant tumors suggest that cancer develops in a multi-step process where randomly accumulated mutations either enhance the expression of proto-oncogenes or

inactivate tumor suppressor genes (Vogelstein, B. and Kinzler, K.W., *Trends Genet.* 9:138-141 (1993); Nowell, P.C., *Science* 194:23-28 (1976)). Frequent loss of heterozygosity in tumor cells points to the possible involvement of at least a dozen recessive mutations in breast cancer (Callahan, R., *et al.*, *Annals N.Y. Acad. Sci.* 698:21-30 (1993)). Genome instability continues to be an important factor even in malignant cells, causing some of them to acquire new alterations that lead to changes such as drug resistance (Cifone, M.A. and Fidler, I.J., *Proc. Natl. Acad. Sci. (USA)* 78:6949-6952 (1981); Horii, A., *et al.*, *Cancer Res.* 54:3373-3375 (1994); Loeb, L.A., *Cancer Res.* 51:3075-3079 (1991); Tlsty, T.D., *et al.*, *Proc. Natl. Acad. Sci. (USA)* 86:9441-9445 (1989)). An important issue in tumor progression is the source of the genome instability. Considerable evidence shows that somatic mutations account for defects in the regulation of proto-oncogenes, and the inactivation of tumor suppressor genes and DNA repair genes. These changes can interact to accelerate genome instability through the mutagenesis that arises from the loss of apoptosis and the inability to control cell division in the presence of persistent DNA damage. Recent evidence shows that aberrant DNA methylation events at CpG sequences provide an alternative route to the inactivation of tumor suppressor genes by inactivating their promoters.

#### CpG island methylation and gene expression.

[0341] DNA methylation in the human genome is most frequent on Cs in the dinucleotide sequence CpG. Methylation at these sites is thought to play roles in a number of processes including gene imprinting, the inactivation of transposable elements, and the inactivation of the X-chromosome in females. In spite of the apparent functional importance of CpG sequences, the human genome has a 4-5 fold lower frequency of CpG dinucleotides than expected given the overall frequency of C and G in human DNA. This distribution probably reflects the gradual selective removal of methylated C driven by the deamination of 5-methylcytosine to thymine. The CpG content of the genome

is organized into two classes. A large fraction of CpG sequences are distributed into about 45,000 clusters known as CpG islands which contain the expected frequency of CpG given the local C + G content. CpG islands range from about 300 to 3000 bp in length and overlap with about 50% of all human promoters (Antequera, F. and Bird, A., *Proc. Natl. Acad. Sci. (USA)* 90:11995-11999 (1993)). In normal cells CpG islands are un-methylated while unclustered CpG sequences outside of CpG islands are uniformly methylated. There is an inverse relationship between the methylation of Cs in CpG sequences and the activities of overlapping promoters. Gradual age related-methylation of CpG islands could lead to a gradual reduction in the expression of affected genes.

**CpG island methylation as a biomarker for carcinogenesis.**

[0342] CpG methylation is potentially a powerful marker for cancer progression. Examination of the promoters of tumor suppressor genes from tumor biopsies suggests that CpG methylation is common enough to equal the impact of mutagenesis in tumor promotion (Toyota, M. and Issa, J.-P., *Electrophoresis* 21:329-333 (2000)). One half of the tumor suppressor genes that are associated with heritable cancers potentially can be inactivated through hypermethylation of their promoters. Although the genomic pattern of CpG methylation is stable over many cell divisions, age related increases in the methylation of CpG islands occur in normal cells. In the case of the ER and Veriscan genes, the age related increases in methylation correlate with reduced gene expression in otherwise normal cells (Issa, J.P., *et al.*, *Nat. Genet.* 7:536-540 (1994); Toyota, M. and Issa, J.P., *Semin. Cancer Biol.* 9:349-357 (1999)). Coupled with the age-related increases in CpG island methylation, the prevalence of susceptible tumor suppressor genes suggests a plausible mechanism for the link between aging and cancer. Age related CpG methylation could contribute the to hyperproliferative state that precedes tumorogenesis as tumor suppressor gene expression is reduced.

- 140 -

[0343] CpG hypermethylation of tumor suppressor gene promoters is an important alternative to mutagenesis in tumorogenesis as demonstrated in comparisons of inherited cancers versus spontaneous cancers. Germline mutations in hMLH1 result in familial colon cancer associated with global alterations in microsatellite repeat sequences due to a defect in mismatch repair. About 10-15% of sporadic colon cancers show microsatellite instability but mutations in the mismatch repair genes are uncommon in non-familial cancers. Up to 80% of sporadic tumors with a mismatch repair defect suffer hypermethylation of the hMLH1 promoter which blocks transcription of the otherwise normal hMLH1 gene (Herman, J.G., *et al.*, *Proc. Natl. Acad. Sci. (USA)* 95:6870-6875 (1998)). hMLH1 expression in tumor cells was restored following reversal of the hypermethylation (Herman, J.G., *et al.*, *Proc. Natl. Acad. Sci. (USA)* 95:6870-6875 (1998)). A similar situation has been documented for breast cancer. Familial breast cancer is strongly associated with mutations in BRCA1. On the other hand, sporadic breast cancer does not involve BRCA1 mutations but is promoted by BRCA1 promoter hypermethylation (Rice, J.C., *et al.*, *Oncogene*. 17:1807-1812 (1998); Esteller, M., *et al.*, *J. Natl. Cancer Inst.* 92:564-569 (2000)). In some cases the mode of inactivation is apparently tissue specific. For example, the p16INK4a/cyclinD-Rb signaling pathway is inactivated in virtually all cancers either by p16INK4a defects or by mutations in the Rb gene. p16INK4a loss occurs by deletion, point mutation or by hypermethylation of the promoter. The mode of inactivation depends on the tumor type. Colon cancers with p16INK4a defects are virtually always associated with p16INK4a promoter hypermethylation (Herman, J., *et al.*, *Cancer Res.* 55:4525-4530 (1995)).

CpG methylation is associated with many cancers and frequently is an early marker for tumorogenesis

[0344] Published surveys of primary tumors from virtually all of the common cancers identified as many as 60 genes that are hypermethylated in sporadic cancers. These genes are involved in all of the physiologically important

aspects of tumorogenesis including tumor suppression, DNA repair, cell adhesion, and apoptosis. CpG methylation patterns are frequently biased to particular genes in particular types of cancers. Therefore, it should be possible to develop methylation signatures for common cancers, indicating both cancer type and stage (Robertson, K., *Oncogene* 20:3139-3155 (2001)). Data on the methylation status of multiple promoters could give clues as to the location of a tumor in cases where several organs can contribute to a sample. For example, shed bladder, kidney or prostate cells can populate a urine sample. Tumors from each of these tissues are frequently associated with distinct combinations of CpG island methylation (FIG. 14).

[0345] In virtually all cases, whether caused by mutagenesis or hypermethylation, defective expression of tumor suppressor genes begins at an early stage in tumor progression. Detection of these early methylation events before advanced symptoms appear should improve the chances that a cancer will be detected while it is highly curable. In one study detection of CpG methylation of the MGMT and p16INK4a promoters in sputum samples from smokers preceded clinical diagnosis of lung cancer by up to 3 years (Palmisano, W.A., *et al.*, *Cancer Res.* 60:5954-5458 (2000)).

[0346] In an exemplary embodiment, the methods of the invention may be utilized to monitor disease initiation, progression, metastasis, recurrence, and any responses to treatment therapies by providing diagnostic techniques, which can detect altered methylation states and patterns. Methylated cytosine residues in a DNA fragment can be detected based upon the resistance of such residues to deamination by a deaminating agent, such as sodium bisulfite for example. When denatured (*i.e.*, single-stranded) DNA is exposed to a deaminating agent, such as sodium bisulfite, unmethylated cytosine (C) residues are converted into uracil residues (U), while methylated cytosine residues (5-mCyt) remain unchanged. That is, as illustrated in FIG. 12, deamination resulting from a treatment with sodium bisulfite causes the originally unmethylated cytosines to change their complementary base-pairing partner from guanine (G) to adenine (A). However, the methylated

cytosines (5-mCyt) retain their base-pairing specificity for G. Thus, after deamination by sodium bisulfite, a target DNA sequence will have only as many complementary CpG islands as there were methylated CpG islands in the original, untreated target DNA sequence. Additionally, as further illustrated in FIG. 12, if an original, untreated target DNA sequence has no methylated CpG islands, then the bisulfite-treated target DNA sequence will no longer contain any CpG islands.

[0347] In view of the foregoing, the level of methylation of the CpG islands in a target DNA sequence may be determined by measuring the relative level of unaltered CpG sites. This relative measurement may be accomplished by initiating abortive transcription at the CpG sites that remain after the target DNA sequence has been exposed to a deaminating agent, such as sodium bisulfite. The sodium bisulfite reaction is performed according to standard techniques. See, e.g., Gonzalgo et al., U.S. Pat. No. 6,251,594. In one embodiment, as illustrated in FIG. 13, a sodium bisulfite-treated DNA target sequence can be incubated with an RNA polymerase and an initiator, such as a mononucleotide initiator (5' R<sub>1</sub>-C-OH 3') for example. The initiator associates with the polymerase and initiates transcription and RNA synthesis at an intact CpG site on the DNA template. Each CpG site can direct the extension of an initiator to synthesize an abortive transcript (e.g., 5' R<sub>1</sub>-CpG-R<sub>2</sub> 3') through the incorporation of a suitable chain terminator (e.g., pppG-R<sub>2</sub>), as illustrated at Sites 1, 3, and 4 in FIG. 13. Either or both of the initiator and a chain terminating nucleotide may be modified with a label moiety (e.g., R<sub>1</sub> and R<sub>2</sub>, respectively) to allow signal detection. In an exemplary embodiment, the transcripts may be detected through fluorescence resonance energy transfer (FRET) for example, as described in detail above (e.g., the primer contains an energy donor (R<sub>1</sub>) at its 5'-end, and the NTP contains an energy acceptor (R<sub>2</sub>) attached to the nucleobase).

[0348] In an alternate embodiment, a sodium bisulfite-treated DNA target sequence may be incubated with an RNA polymerase and a dinucleotide initiator (5' R<sub>1</sub>-CpG-OH 3'). The initiator then associates with the polymerase

and initiates transcription and RNA synthesis at an intact CpG site on the DNA template. Each CpG site then directs the extension of the dinucleotide initiator to synthesize an abortive trinucleotide transcript through the incorporation of a suitable chain terminator. The nucleotide analog that comprises the chain terminator will depend upon the DNA template sequence. For example, at Site 1 of FIG. 13, a suitable chain terminator would include 5' pppA-R<sub>2</sub> 3', and the resultant abortive trinucleotide transcript would be 5' R<sub>1</sub>-CpGpA-R<sub>2</sub> 3'.

[0349] In another embodiment, as diagrammatically illustrated in FIG. 13, after the target DNA sequence has been deaminated, such as by treating the target DNA sequence with sodium bisulfite for example, a target site probe may be used to form a bubble complex that comprises a target CpG site on the target DNA sequence. In this embodiment, the target site probe is used to direct the RNA polymerase to the target CpG site by positioning the target CpG site at the junction of a single-stranded bubble region and a downstream duplex region on the target DNA sequence. In the illustrated embodiment, the target site probe comprises about 18-54 nucleotides: a first region which hybridizes to the target DNA sequence upstream of the target site comprises about 5-20 nucleotides; an internal second region of non-base-paired nucleotides comprises about 8-14 nucleotides; and a third region which hybridizes to the target DNA sequence downstream of the target site comprises about 5-20 nucleotides. The target site probe may be hybridized to the target DNA sequence either before or while the DNA target sequence is incubated with an RNA polymerase and a suitable RNA initiator. The polymerase associates with the RNA initiator and initiates transcription and RNA synthesis at the CpG site on the DNA template. The polymerase extends the initiator to synthesize an abortive oligonucleotide transcript through the incorporation of a suitable chain terminator. Either or both of the initiator and a chain terminating nucleotide may be modified with a label moiety to allow signal detection, such as by fluorescence resonance energy transfer for example, as described in detail above.

[0350] In another embodiment, methylation of CpG sites may be assessed without deamination.

[0351] In another embodiment, capture probes may be designed to capture the genes of interest, and abortive transcription initiation used to determine the methylation status of the desired genes. For example, genes known to be associated with the progression of a particular cancer, such as colon cancer, may be monitored, including but not limited to APC (adenomatous polyposis coli), CALCA (calcitonin), ER (estrogen receptor), GSTP1, HIC1 (hypermethylated in cancer-1), hMLH1, HPP1/TR/TENB2/TMEFF2 (Transmembrane protein with EFG-like and two follistatin-like domains 2), LKB1/STK11, IGF2 IGF2 (Insulin-like growth factor), MGMT (O<sup>6</sup> methyl guanine methyl transferase 1), MINT25, p14(ARF), p16(INK4a)/MTSI /CDKN2A, PAX6 (paired box gene 6), RAR-Beta2, THBS1 (thrombospondin-1), Veriscan, and WT1 (Wilm's tumor suppressor). Each gene of interest could be removed from the sample by hybridization to a capture sequence, which is unique for the gene of interest. The capture sequence may be immobilized on a solid matrix, including but not limited to magnetic beads, microtiter plates, sepharose, agarose, cation exchange resins, lateral flow strips, glass beads, and microarray chips. Once the gene of interest has been removed from the sample, abortive transcription initiation can be used to determine the methylation status for each particular gene.

[0352] An illustrative procedure for detecting DNA methylation states and patterns, therefore, may include: (a) optionally immobilizing an oligonucleotide capture probe which is specific for a region near a CpG island of a target gene; (b) optionally treating the oligonucleotide capture probe with a denatured DNA sample which potentially contains a target DNA sequence; (c) converting any unmethylated cytosine residues on the target DNA sequence to uracil residues and leaving any methylated cytosine residues unaltered; (d) optionally hybridizing the target DNA sequence with a target site probe; (e) modifying at least one of an RNA initiator and nucleotides comprising a chain terminator to enable detection of the oligonucleotide

transcript; (f) hybridizing the target DNA with the RNA initiator; and (g) extending the RNA initiator with an RNA polymerase such that the RNA polymerase reiteratively synthesizes an oligonucleotide transcript that is complementary to a target site by incorporating complementary nucleotides comprising a chain terminator and releasing an abortive oligonucleotide transcript without either translocating from an enzyme binding site or dissociating from the target DNA sequence; and (g) detecting and optionally quantifying the multiple abortive oligonucleotide transcripts.

[0353] In another embodiment CpG methylation is assessed in two assays. The first assay will reveal the overall level of methylation of a CpG island by measuring methylation of multiple specific CpG sites in a single reaction. If the results of this assay indicate that the CpG island is methylated to an abnormal level, then a second high-throughput methylation assay is performed to assess the degree of methylation at individual CpG sites.

[0354] The CpG epigenetic assay system is designed to discriminate between cytosine and 5-methylcytosine based on sequence changes created by exposure of patient DNA samples to sodium bisulfite. The bisulfite treatment is preceded by the immobilization of the patient DNA. This approach facilitates buffer exchange without losses from the small DNA samples (<1 µg of DNA). FIG. 29 depicts preparation for the CpG assay. Patient DNA strands are separated and affinity tagged by primer extension using biotin labeled primers that flank the CpG island (FIG. 29, A). The hybrid duplexes are immobilized on a covalently-linked streptavidin plate where the 5'-ends of the patient DNA strands are labeled with thiophosphates (FIG. 29, C and D). The patient strands then are separated from the biotinylated DNA strands and covalently linked via their thiophosphates to a maleimide plate. At this stage the DNA is ready for the sodium bisulfite-mediated deamination reaction.

[0355] The most commonly used bisulfite treatment method for measuring CpG methylation causes considerable DNA damage due to lengthy incubation of free DNA strands at high temperature and low pH (typically 16 hours at 55 °C and pH 5) (Frommer, M., *et al.*, *Proc. Natl. Acad. Sci. (USA)* 89:1827-1831

(1992)). High temperature incubation is required in most conventional DNA methylation assays to prevent hairpin formation which would protect segments of the DNA from deamination. It is possible to carryout this reaction under relatively mild conditions (35 °C for 4 hr at pH 5) because, it is believed, the targeted CpGs remain single-stranded through their association with a deoxynucleotide target site probe (TSP). The TSPs are used in the final methylation assay to target specific GpCs by placing them in partially duplex structures (FIG. 29, F and H). Multiple TSPs are annealed to the island DNA in the CpG-detection assay to get an aggregate measure of CpG methylation. In the second CpG assay a single TSP is annealed per well.

[0356] The discrimination between methylated and un-methylated sites is performed by an RNA polymerase which reiteratively produces short oligonucleotide products from CpG sites. Individual CpGs are targeted by 2 components: the oligo-deoxynucleotide target site probe (TSP) and a di- or trinucleotide initiator. The initiator CpG is specific for methylated CpG sites and the initiator CpA is specific for deaminated sites which have been converted to UpG by bisulfite treatment. Alignment of the appropriate initiator allows the incorporation of a radioactive NTP encoded by the base adjacent to the CpG. The oligonucleotide products of CpGpN and CpApN can be separated by thin layer chromatography and detected by autoradiography. Both initiators can be included in the same reaction to give a methylation index defined as pmoles CpGpN/(pmoles CpGpN + CpApN).

[0357] Little information exists about the role of CpG methylation in skin cancers other than melanoma. Consequently, the present invention is useful in determining the methylation state of cancer and, as such, in the development of improved diagnostic tests for a wide range of cancers.

#### Genetic Mutations

[0358] In another aspect of the invention, the methods disclosed herein may be used in diagnostic assays which detect mutations in the form of gross

chromosomal rearrangements or single or multiple nucleotide alterations, substitutions, insertions, or deletions. In an exemplary embodiment, as diagrammatically illustrated in FIG. 15, single nucleotide polymorphisms (SNPs) may be detected through the use of an abortive oligonucleotide synthesis reaction. A known target SNP sequence (e.g., 3' dN<sub>X</sub>pdN<sub>Y</sub>pdN<sub>T</sub> 5', where dN<sub>T</sub> is a target SNP site) can be incubated with an RNA polymerase, an RNA initiator, such as a dinucleotide initiator for example, and nucleotides (e.g., a chain terminator such as 5' pppN<sub>T</sub>-R<sub>2</sub>). The initiator binds immediately upstream of the target SNP sequence, associates with the polymerase, and initiates transcription and RNA synthesis at the target SNP site. In one embodiment, the polymerase elongates the initiator by incorporating the chain terminator to produce an abortive trinucleotide product. Either or both of the initiator and a chain terminating nucleotide may be modified with a label moiety (R<sub>1</sub> and R<sub>2</sub>, respectively) to allow signal detection. In an exemplary embodiment, the transcripts may be detected through fluorescence resonance energy transfer (FRET) for example, as described in detail above (e.g., the initiator contains an energy donor (R<sub>1</sub>) at its 5'-end, and the chain terminator contains an energy acceptor (R<sub>2</sub>) attached to the nucleobase).

[0359] An illustrative procedure for detecting mutations in a target DNA sequence (FIG. 16), therefore, may include the following: (a) optionally immobilizing a capture probe designed to hybridize with a target DNA sequence which includes a mutation; (b) optionally hybridizing the capture probe with a DNA sample which potentially contains the target DNA sequence; (c) optionally hybridizing the target DNA sequence with a target site probe; (d) modifying at least one of an RNA initiator (R<sub>1</sub>N<sub>I</sub>-OH) and nucleotides comprising a chain terminator (pppN<sub>T</sub>-R<sub>2</sub>) to enable detection of the oligonucleotide transcript synthesized by the RNA polymerase; (e) hybridizing the target DNA sequence with the RNA initiator; (f) extending the RNA initiator with an RNA polymerase such that the RNA polymerase reiteratively synthesizes an oligonucleotide transcript that is complementary to

a target mutation site by incorporating complementary nucleotides comprising a chain terminator and releasing an abortive oligonucleotide transcript without either translocating from an enzyme binding site or dissociating from the target DNA sequence; and (g) detecting and optionally quantifying the multiple abortive oligonucleotide transcripts.

#### Pathogenic Organisms

[0360] In another aspect of the invention, the methods disclosed herein may be used in diagnostic assays which detect the presence of a particular nucleic acid (DNA or RNA), thereby serving to indicate the presence of either a particular or a generic organism which contains the gene, or which permit genetic typing of a particular organism without the need for culturing the organism. The test sample may be suspected of containing a target nucleic acid sequence from a particular microorganism, such as bacteria, yeast, viruses, viroids, molds, fungi, and the like. The test sample may be collected from a variety of sources including but not limited to, animal, plant or human tissue, blood, saliva, semen, urine, sera, cerebral or spinal fluid, pleural fluid, lymph, sputum, fluid from breast lavage, mucosal secretions, animal solids, stool, cultures of microorganisms, liquid and solid food and feedproducts, waste, cosmetics, air, and water.

[0361] In another aspect of the invention, the methods disclosed herein may be used in diagnostic assays which detect the presence of a particular nucleic acid (DNA or RNA), thereby serving to indicate the presence of either a particular or a generic pathogenic organism which contains the gene, or which permit genetic typing of a particular organism without the need for culturing the organism. In an exemplary embodiment, as diagrammatically illustrated in FIG. 17, an oligonucleotide capture probe that is sequence-specific for a target pathogen polynucleotide is attached to a solid matrix, such as a microtiter plate for example, and the capture probe is treated under hybridizing conditions with a test sample which potentially contains the target pathogen

polynucleotide. The test sample may be suspected of containing a target nucleic acid sequence from a particular pathogen, such as, for example, a microorganism, such as bacteria, yeast, viruses, viroids, molds, fungi, and the like. The test sample may be collected from a variety of sources including but not limited to, animal, plant or human tissue, blood, saliva, semen, urine, sera, cerebral or spinal fluid, pleural fluid, lymph, sputum, fluid from breast lavage, mucosoal secretions, animal solids, stool, cultures of microorganisms, liquid and solid food and feedproducts, waste, cosmetics, air, and water.

[0362] The target pathogen polynucleotide may be either RNA or DNA. A target pathogen polynucleotide that is present in the test sample hybridizes to the capture probe, and washing is then performed to remove any components of the test sample that were not immobilized by the capture probe. Target DNA or RNA may be retrieved by addition of specific sequences via primer extension, for example. In an exemplary embodiment, the captured target pathogen polynucleotide is hybridized with an artificial promoter cassette (APC). The APC linker sequence includes a single-stranded overhang region on either its 3' or 5' end (depending upon the orientation needed to create an antiparallel hybrid with the capture probe). In other words, the APC linker is complementary to the sequence on the free end of the captured target pathogen polynucleotide, thereby permitting the APC linker to hybridize to the target pathogen polynucleotide.

[0363] An initiator and a polymerase are added to the reaction mixture. The initiator hybridizes within the bubble region of the APC at a position that facilitates catalysis of a synthesis reaction by a suitable polymerase at the target site. The initiator may be RNA or DNA, may comprise from about 1 to 25 nucleotides, and may include one or more nucleotide analogs as well as nucleotides. The polymerase may be an RNA-dependent or DNA-dependent RNA polymerase. The DNA or RNA APC may or may not be attached to other molecules, such as proteins, for example. In an exemplary embodiment, the APC comprises DNA, the initiator is RNA, and the polymerase is a DNA-dependent RNA polymerase.

[0364] During the polymerization reaction, the initiator is extended or elongated by the polymerase through the incorporation of nucleotides that have been added to the reaction mixture. As the polymerase reaction proceeds, the polymerase extends the initiator, as directed by the APC template sequence within the bubble region, by incorporating complementary nucleotides, including a suitable chain terminator, that are present in the reaction mixture. When the polymerase incorporates a chain terminator into the nascent oligonucleotide product, chain elongation terminates due to the polymerase's inability to catalyze the addition of a nucleotide at the 3' position on the pentose ring of the incorporated chain terminator. Consequently, the polymerase aborts the initiated synthesis event by releasing the oligonucleotide product and reinitiating the synthesis reaction at the target site. Either or both of the initiator and a chain terminating nucleotide may be modified with a label moiety to allow signal detection. In an exemplary embodiment, the oligonucleotide products may be detected through fluorescence resonance energy transfer (FRET), as described above (e.g., the initiator contains an energy donor ( $R_1$ ) at its 5'-end, and the chain terminator contains an energy acceptor ( $R_2$ ) attached to the nucleobase).

[0365] An illustrative procedure for detecting the presence of pathogens (FIG. 18), therefore, may include: (a) optionally immobilizing a capture probe designed to hybridize with a target pathogen polynucleotide; (b) optionally hybridizing the capture probe with a test sample which potentially contains a target pathogen polynucleotide. The target nucleic acid may be copied to DNA via reverse transcription (for RNA pathogens) or primer extension (for DNA pathogens). In both bases, a DNA sequence corresponding to the Artificial promoter cassette (APC) linker will be added to the target copy (FIG. 1); (c) optionally washing the captured target pathogen polynucleotide to remove any unhybridized components of the test sample; (d) hybridizing the captured target pathogen polynucleotide with an abortive promoter cassette; (e) modifying at least one of a initiator and nucleotides comprising a chain terminator to enable detection of the oligonucleotide product synthesized by

the polymerase; (f) hybridizing the artificial promoter cassette with a initiator; (g) extending the initiator with a polymerase such that the polymerase reiteratively synthesizes an oligonucleotide product that is complementary to a target site by incorporating complementary nucleotides comprising a chain terminator and releasing an abortive oligonucleotide product without either translocating from an enzyme binding site or dissociating from the APC; and (h) detecting and optionally quantifying the multiple abortive oligonucleotide products.

[0366] The present invention is useful for detecting pathogens in mammals. In particular the invention is useful for the detection of bacteria, viruses, fungus, molds, amoebas, prokaryotes, and eukaryotes. Preferred mammals include monkeys, apes, cats, dogs, cows, pigs, horses, rabbits and humans. Particularly preferred are humans.

[0367] The methods of the invention are particularly useful for monitoring the presence or absence of pathogenic nucleic acids and proteins. The invention can be used to detect, diagnose, and monitor diseases, and/or disorders associated with pathogenic polypeptides or polynucleotides. The invention provides for the detection of the aberrant expression of a polypeptide or polynucleotide. The method comprises (a) assaying the expression of the polypeptide or polynucleotide of interest in cells, tissue or body fluid of an individual using the methods of abortive initiation transcription described above, and (b) comparing the level of gene expression, protein expression, or presence of sequences of interest with a standard gene or protein expression level or sequence of interest, whereby an increase or decrease in the assayed polypeptide or polynucleotide level compared to the standard level is indicative of aberrant expression indicating presence of a pathogen of interest..

[0368] The presence of an abnormal amount of transcript in biopsied tissue or body fluid from an individual may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative

measures or aggressive treatment earlier thereby preventing the development or further progression of the disease caused by the pathogen.

[0369] The invention is particularly useful for monitoring the presence of pathogenic organisms including but not limited to *E. coli*, *Steptococcus*, *Bacillus*, *Mycobacterium*, HIV, and Hepatitis.

[0370] The methods of the invention may be used to test for pathogenic microorganisms in aqueous fluids, in particular water (such as drinking water or swimming or bathing water), or other aqueous solutions (such as fermentation broths and solutions used in cell culture), or gases and mixtures of gases such as breathable air, and gases used to sparge, purge, or remove particulate matter from surfaces. Breathable air from any source including but not limited to homes, schools, classrooms, workplaces, aircraft, spacecraft, cars, trains, buses, and any other building or structure where people gather, may be tested for the presence of pathogenic microorganisms.

#### mRNA Expression

[0371] In another aspect of the invention, the methods disclosed herein may be used in diagnostic assays which detect messenger RNA (mRNA) expression levels in a quantitative or non-quantitative manner. In an exemplary embodiment, as diagrammatically illustrated in FIG. 19, an oligonucleotide capture probe that is sequence-specific for a target mRNA sequence is attached to a solid matrix, such as a microtiter plate for example, and the capture probe is treated under hybridizing conditions with a test sample which is suspected of containing the target mRNA sequence. A target mRNA sequence that is present in the test sample hybridizes to the capture probe, and washing is then performed to remove any components of the test sample that were not immobilized by the capture probe. The captured target mRNA sequence is then hybridized with an artificial promoter cassette (APC). In the illustrated embodiment, the APC has an APC linker sequence which includes a single-stranded poly-T overhang on its 3' end that is

complementary to the poly-A tail on the 3' end of the target mRNA sequence, thereby permitting the APC linker to hybridize to the poly-A tail of the target mRNA.

[0372] An initiator and a polymerase are added to the reaction mixture. The initiator hybridizes within the bubble region of the APC, upstream of the target site, and facilitates catalysis of a synthesis reaction by a suitable polymerase at the target site. The initiator may comprise from about 1 to 25 nucleotides, and may include one or more nucleotide analogs as well as nucleotides. The polymerase may be an RNA-dependent or DNA-dependent RNA polymerase. The APC may or may not be attached to other molecules, such as proteins, for example. In an exemplary embodiment, the APC comprises DNA, the initiator is RNA, and the polymerase is a DNA-dependent RNA polymerase.

[0373] During the polymerization reaction, the initiator is extended or elongated by the polymerase through the incorporation of nucleotides which have been added to the reaction mixture. As the polymerase reaction proceeds, the polymerase extends the initiator, as directed by the APC template sequence within the bubble region, by incorporating complementary nucleotides, including a chain terminator, that are present in the reaction mixture. When the polymerase incorporates a chain terminator into the nascent oligonucleotide product, chain elongation terminates due to the polymerase's inability to catalyze the addition of a nucleotide at the 3' position on the pentose ring of the incorporated chain terminator. Consequently, the polymerase aborts the initiated synthesis event by releasing the oligonucleotide product and reinitiating the synthesis reaction at the target site. Either or both of the initiator and a chain terminating nucleotide may be modified with a label moiety to allow signal detection, such as by fluorescence resonance energy transfer for example, as described in detail above.

[0374] An illustrative procedure for detecting mRNA expression levels, therefore, may include: (a) optionally immobilizing a capture probe designed to hybridize with a specific or general mRNA sequence; (b) optionally

hybridizing the capture probe with a test sample which potentially contains a target mRNA sequence; (c) optionally washing the captured target mRNA sequence to remove any unhybridized components of the test sample; (d) hybridizing the captured target mRNA sequence with an abortive promoter cassette; (e) modifying at least one of a initiator and nucleotides comprising a chain terminator to enable detection of the oligonucleotide product synthesized by the polymerase; (f) hybridizing the artificial promoter cassette with the initiator; (g) extending the initiator with a polymerase such that the polymerase reiteratively synthesizes an oligonucleotide product that is complementary to a target site by incorporating complementary nucleotides comprising a chain terminator and releasing an abortive oligonucleotide product without either translocating from an enzyme binding site or dissociating from the APC; and (h) detecting and optionally quantifying the multiple abortive oligonucleotide products.

#### Protein Detection

[0375] In another aspect of the invention, the methods disclosed herein may be used in diagnostic assays which detect proteins. As shown in Figure 20, an artificial promoter cassette linker can be made with a protein modifier group attached, such that the linker is complementary to the APC linker attached to the APC.

[0376] An illustrative procedure for detecting proteins, therefore, may include: (a) attaching a short piece of DNA of a defined sequence (APC linker) to a protein via a primary amine, a secondary amine, or a sulphydral group; (b) retrieving and immobilizing the modified protein with an antibody or some other affinity agent against the protein; and (c) attaching an artificial promoter cassette to the protein by hybridization of the APC cassette to the APC linker on the labeled protein; (d) detecting the protein by (i) treating the DNA with an initiator nucleotide under hybridizing conditions; and (ii) treating the DNA

with an RNA polymerase and nucleotides or nucleotide analogs that permit detection. Process (d) occurs repeatedly for each RNA polymerase bound.

#### Cancer Detection

[0377] The present invention is useful for detecting cancer in mammals. In particular the invention is useful during diagnosis of cancer. Preferred mammals include monkeys, apes, cats, dogs, cows, pigs, horses, rabbits and humans. Particularly preferred are humans.

[0378] The methods of the invention are particularly useful for monitoring the status of DNA methylation, genetic mutations, mRNA expression patterns, and protein expression patterns. The invention can be used to detect, diagnose, and monitor diseases, and/or disorders associated with the aberrant expression and/or activity of a polypeptide or polynucleotide. The invention provides for the detection of the aberrant expression of a polypeptide or polynucleotide, the presence of mutations, and changes in methylation status of DNA. The method comprises (a) assaying the expression of the polypeptide or polynucleotide of interest in cells, tissue or body fluid of an individual using the methods of abortive initiation transcription described above, and (b) comparing the level of gene expression, protein expression, or presence of sequences of interest with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide or polynucleotide level compared to the standard level is indicative of aberrant expression indicating presence of cancer or a pathogen of interest.

[0379] The presence of an abnormal amount of transcript in biopsied tissue or body fluid from an individual may indicate a predisposition for the development of cancer or a disease of interest, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby

preventing the development or further progression of the cancer or disease caused by the pathogen.

[0380] The diagnostic assays of the invention can be used for the diagnosis and prognosis of any disease, including but not limited to Alzheimer disease, muscular dystrophy, cancer, breast cancer, colon cancer, cystic fibrosis, fragile X syndrome, hemophilia A and B, Kennedy disease, ovarian cancer, lung cancer, prostate cancer, retinoblastoma, myotonic dystrophy, Tay Sachs disease, Wilson disease, and Williams disease. These assays are believed to be particularly useful for the diagnosis and prognosis of all types of cancer.

#### Kits of the Invention

[0381] The invention also provides kits for carrying out the methods of the invention. Such kits comprise, in one or more containers, usually conveniently packaged to facilitate their use in assays, quantities of various compositions essential for carrying out the assays in accordance with the invention. Thus, the kits comprise one or more initiators according to the invention. The kits may additionally comprise an enzyme with polymerase activity, such as an RNA and/or DNA polymerase for example, to extend the primer of the kit, as well as reagents for processing a target nucleic acid. The kit may also comprise nucleotides and/or nucleotide analogs to enable detection of the oligonucleotide products synthesized by the methods of the invention. The kits may also include oligonucleotide target site probes for forming a bubble complex on the target nucleic acid. The kit may also contain an abortive promoter cassette. The kits may also contain components for the collection and transport of materials, including but not limited to, membranes, affinity materials, test tubes, petri dishes, and dipsticks. The kit may also include microtiter plates, bio-chips, magnetic beads, gel matrices, or other forms of solid matrices to which an oligonucleotide capture probe, which is specific for a particular target sequence, has been bound. The relative amounts of the components in the kits can be varied to provide for reagent

- 157 -

concentrations that substantially optimize the reactions involved in the practice of the methods disclosed herein and/or to further optimize the sensitivity of any assay.

[0382] The test kits of the invention can also include, as is well-known to those skilled in the art, various controls and standards, such as solutions of known target nucleic acid concentration, including no target sequence (negative control), to ensure the reliability and accuracy of the assays carried out using the kits and to permit quantitative analyses of test samples using the kits. The kits include a set of instructions, which are generally written instructions, though the instructions may be stored on electronic storage media (e.g., magnetic diskette or optical disk), relating to the use of the components of the methods of the invention. The instructions provided with the kit generally also include information regarding reagents (whether included or not in the kit) necessary or preferred for practicing the methods of the invention, instructions on how to use the kit, and/or appropriate reaction conditions.

#### EXAMPLES

[0383] The following examples are provided for purposes of illustration only and not of limitation. Those of skill in the art will readily recognize a variety of non-critical parameters which could be changed or modified to yield essentially similar results.

#### EXAMPLE 1

##### RNA Primer-Initiated Abortive Transcription with an RNA Polymerase

[0384] Reaction conditions have been optimized for abortive transcription initiation. The components and concentrations of Buffer T favor abortive transcription initiation. Buffer T is comprised of: 20 mM Tris-HCl pH 7.9, 5 mM MgCl<sub>2</sub>, 5 mM beta-mercaptoethanol, 2.8% (v/v) glycerol. Primers are either ribonucleoside-triphosphates (NTPs) or dinucleotides ranging in concentration from 0.2-1.3 mM. Final NTP concentrations range from 0.2-1.3

mM. The high ends of the concentration ranges are designed for preparative abortive transcription. The template DNA concentration is less than 2  $\mu$ M in terms of phosphate. *E. coli* RNA polymerase is added to a final concentration of between 15 nM and 400 nM. Either holoenzyme or core can be used with a single-stranded template DNA. Yeast inorganic pyrophosphatase is added to 1 unit/ml in preparative reactions to prevent the accumulation of pyrophosphate. At high concentrations pyrophosphate can reverse the synthesis reaction causing RNA polymerase to regenerate NTPs at the expense of the RNA products. One unit of pyrophosphatase is defined as the amount of enzyme to liberate 1.0  $\mu$ M of inorganic orthophosphate per min. at 25 °C and pH 7.2. Reactions are incubated at 37°C for up to 72 hours for preparative reactions. These conditions are representative; for specific templates, optimization of particular components and concentrations may enhance the efficiency of abortive initiation.

[0385] Three different initiators were used in this example: (1) TAMARA-ApG; (2) Biotin ApG; and (3) ApG. The target nucleic acid template was denatured by boiling for 5 minutes at 95°C and immediately placing on ice. Each reaction was prepared as follows:

5.0  $\mu$ l 1X Buffer T  
2.5  $\mu$ l of a-32P-UTP  
14.3  $\mu$ l ddH2O  
1  $\mu$ l of *E. coli* RNA polymerase (1U/ $\mu$ l)  
100 ng (2  $\mu$ l) of template DNA  
10 nmoles (1.2  $\mu$ l) of initiator  
22.8  $\mu$ l of reaction buffer

[0386] Incubate at 37°C for 12-16 hours. Thin layer chromatography was performed using standard methods known in the art to determine the extent of incorporation of UTP in the third position (Figure 25).

[0387] Both TAMARA-ApG and biotin ApG allowed for incorporation of the nucleotide UTP. Biotin ApG incorporated more efficiently than TAMARA-ApG, but not as efficient as ApG.

#### EXAMPLE 2

##### Abortive Initiation Reaction with a Labeled Terminator

[0388] Abortive transcription initiation reactions may be performed with a labeled initiator and/or a labeled terminator. The following reaction conditions were used to incorporate a labeled terminator:

5  $\mu$ l 1X Buffer T  
3  $\mu$ l 100 ng denatured DNA template (pBR322)  
13.5  $\mu$ l dd H<sub>2</sub>O  
1  $\mu$ l E. coli RNA polymerase  
1.2  $\mu$ l dinucleotide initiator ApG  
1.5  $\mu$ l of 7mM SF-UTP

[0389] Incubate mixture at 37°C for 16 hours in temperature controlled microtitre plate reader. Thin layer chromatography was performed using standard methods known in the art, and demonstrated that the labeled trinucleotide ApGpU was generated (Figure 26).

#### EXAMPLE 3

##### Fluorescent Energy Transfer Between Donors and Acceptors

[0390] The above examples have demonstrated that both labeled initiators and terminators can be incorporated into the oligonucleotide products. One efficient method to measure incorporation of the labeled nucleotides is by Fluorescent Resonance Energy Transfer. The following conditions were used to demonstrate FRET between a labeled initiator and a labeled terminator:

5  $\mu$ l 1X Reaction Buffer (Buffer T)

- 160 -

3  $\mu$ l denatured DNA template (300 ng pBR322)  
13.5  $\mu$ l dd H<sub>2</sub>O  
1  $\mu$ l E.coli RNA polymerase  
1.2  $\mu$ l Initiator (TAMARA-ApG or ApG or Biotin-ApG)  
1.5  $\mu$ l of 7mM SF-UTP

[0391] The reaction mixture was incubated at 37C for 16 hours in temperature controlled microtitre plate reader, which was set to read at the following parameters: Ex 485, Em 620, Gain 35, 99 reads/well/cycle. Under the reaction conditions described above, the RNA polymerase reiteratively synthesizes an oligonucleotide product composed of the initiator (TAMARA-SpApG) and the terminator (SF-UTP).

[0392] Formation of the oligonucleotide product, TAMARA-SpApGpU-SF, places the initiator and the terminator within 80 angstroms of each other, which allows for the transfer of energy between the chemical moieties. Energy is transferred from the donor, which is SF-UTP, to the acceptor, which is TAMARA-ApG. This transfer of energy can be detected and/or quantitated by a change in wavelength emission of TAMARA (TAMARA Absorbance = 540 nm; Emission = 565 nm)

[0393] As the oligonucleotide product is generated, energy transfer occurs between TAMARA-SpApG and SF-UTP, which changes the wavelength at which TAMARA emits. If RNA polymerase or DNA is omitted from the reaction, there is no transfer of energy between the initiator and the terminator, and no change in the wavelength at which TAMARA emits.

[0394]

#### EXAMPLE 4

##### Determination of the Methylation Status of Specific Residues of the CDKN2A gene

[0395] The sample to be analyzed is collected from a human stool sample. Methods of DNA extraction from stool samples are well known in the art, and

commercial kits are available for extracting human DNA from stool samples, such as QIAamp DNA Stool Mini Kit from Qiagen (Valencia, CA).

[0396] After extraction, the sample is applied to the wells of a microtiter plate, which contain a capture probe for the gene of interest, in this particular example, the capture probe is for CDKN2A gene. The nucleotide sequence of a representative capture probe for the CpG islands of the CDKN2A gene is as follows:

[0397] ATATACTGGGTCTACAAGGTTAACGTCAACCAGGGATTGAAATATAACTT TTAAACAGAGCTGG. The DNA sample is incubated with the capture probe to allow hybridization. A representative hybridization protocol is as follows: (1) prehybridize with 2.5X SSC, 5X Denhardts at room temperature for 30 minutes; (2) hybridize with 2.5X SSC, 5X Denhardts, 30% formamide at room temperature for 2 hours; (3) wash twice with 1X SSC at 42°C for 10 minutes, maintaining 42°C; and (4) wash three times with 0.1X SSC at 42°C for 10 minutes, maintaining 42°C.

[0398] The DNA is treated with a deaminating agent, such as sodium bisulfite, which will de-amine the unmethylated C's in the DNA, while leaving the methylated C's unaltered. The wells are then washed under medium stringency conditions to remove the remaining sodium bisulfite.

[0399] A representative transcription reaction is comprised of the following components: *E. coli* holoenzyme RNA polymerase; reaction buffer: 10nM Tris-HCl, pH 7.0; 10 mM KCl; 0.5 mM Na<sub>2</sub>EDTA; and 50 mg/ml BSA; an initiator, and nucleotide analogs. The reaction conditions for particular nucleotide sequence may vary. Other polymerases may be used, such as *E. coli* T7, or SP6. The reaction buffer can be optimized to increase abortive initiation events by adjusting the salt concentration, divalent cations and concentrations, the glycerol content, and the amount and type of reducing agent to be used.

[0400] The initiator will be a 5'-αSpCpG dimer labeled through the 5'-S with fluorescein, which functions as the donor in this reaction. The nucleotide analog(s) will be labeled with TAMARA, which will function as the acceptor

- 162 -

in this reaction. The initiator can be labeled with either the donor or the acceptor in the FRET reaction, and depending upon the fluorescent molecule used to label the initiator, the nucleotide analog(s) will be labeled with either a donor or an acceptor.

[0401] Fluorescein is excited using a 360 nm wavelength filter; the resulting emission peak is at about 515 nm. If the TAMARA is in close proximity to the fluorescein, it becomes excited at 542 nm, resulting in an emission peak of 568 nm. The near ultraviolet wavelength excites the fluorescein but not the rhodamine. Therefore signal will only be generated if the fluorescein is in close proximity to the rhodamine. This signal can be generated and monitored in a fluorescent microtitre plate reader that has been fitted with specific excitation and emission filters for this FRET pair. These filters and plate readers are commercially available from a number of sources, although most clinical labs and research facilities already use a fluorescent microtitre plate reader.

#### EXAMPLE 5

##### Measurement of CpG methylation.

[0402] CpG methylation is assessed in two assays. The first assay determines the overall level of methylation of a CpG island by measuring methylation of multiple specific CpG sites in a single reaction. If this assay indicates that the CpG island is methylated to an abnormal level, a second high-throughput methylation assay is performed on individual CpG sites.

##### Preparation of patient DNA for deamination.

[0403] The CpG epigenetic assay system is designed to discriminate between cytosine and 5-methylcytosine based on sequence changes created by exposure of patient DNA samples to sodium bisulfite. The bisulfite treatment is preceded by the immobilization of patient DNA. This approach facilitates

buffer exchange without losses from the small DNA samples (<1  $\mu$ g of DNA). FIG. 29 depicts the preparation for the CpG detection assay. Patient DNA strands are separated and affinity tagged by primer extension using biotin labeled primers that flank the CpG island (FIG. 29, A). The hybrid duplexes are immobilized on a covalently-linked streptavidin plate where the 5'-ends of the patient DNA strands are labeled with thiophosphates (FIG. 29, C and D). The patient strands are then separated from the biotinylated DNA strands and covalently linked via their thiophosphates to a maleimide plate. At this stage the DNA is ready for the sodium bisulfite-mediated deamination reaction.

[0404] The most commonly used bisulfite treatment method for measuring CpG methylation causes considerable DNA damage due to lengthy incubation of free DNA strands at high temperature and low pH (typically 16 hours at 55 °C and pH 5) (Frommer, M., *et al.*, *Proc. Natl. Acad. Sci. (USA)* 89:1827-1831 (1992)). High temperature incubation is required in most conventional DNA methylation assays to prevent hairpin formation which would protect segments of the DNA from deamination. Under the current method, it is possible to carryout this reaction under relatively mild conditions (35 °C for 4 hr at pH 5) because, it is believed, the targeted CpGs remain single-stranded through their association with a deoxynucleotide target site probe (TSP). The TSPs are used in the final methylation assay to target specific GpCs by placing them in partially duplex structures (FIG. 29, F and H). Multiple TSPs are annealed to the island DNA in the first assay to get an aggregate measure of CpG methylation. In the second assay a single TSP is annealed per well.

#### Measurement of CpG methylation by the RNA polymerase transcription reaction

[0405] The discrimination between methylated and un-methylated sites is performed by an RNA polymerase which reiteratively produces short oligonucleotide products from CpG sites. Individual CpGs are targeted by 2 components: the oligo-deoxynucleotide target site probe (TSP) and a di- or trinucleotide initiator. The initiator CpG is specific for methylated CpG sites

and the initiator CpA is specific for deaminated sites which have been converted to UpG by bisulfite treatment. Alignment of the appropriate initiator allows the incorporation of a radioactive NTP encoded by the base adjacent to the CpG. The oligonucleotide products of CpGpN and CpApN are then separated by thin layer chromatography and detected by autoradiography. Both initiators are included in the same reaction to give a methylation index defined as pmoles CpGpN/pmoles CpGpN + CpApN.

#### EXAMPLE 6

##### The role of CpG islands in cancer

[0406] Little information exists about the role of CpG methylation in skin cancers other than melanoma. Consequently, methylation studies will focus on 6 CpG islands that are aberrantly methylated in a wide range of cancers and one island that is associated with melanoma (Table 2).

TABLE 2  
CpG methylation of cancer.

Gene	Function	Tumor types with CpG methylation
MAGE1	Expressed in testis and melanoma	Melanoma, lung cancer
CDKN2A (p16INK4a)	cyclin-dependent kinase inhibitor	At least 14 tumor types
CDKN1B (p27KIP1)	cyclin-dependent kinase inhibitor	Melanoma
MGMT	O6-methylguanine-DNA methyltransferase. Hypermethylation of the gene is associated with drug resistance in melanoma	Brain, Colon, Lung, Breast, Esophageal adenocarcinoma.
COX2	Prostaglandin-endoperoxide synthase 2	Breast, Prostate, Colon
RAR-beta2	Retinoic acid receptor	Colon, Breast, Lung, Pancreatic
RASSF1A	Interferes with accumulation of cyclin D1	Breast, Ovarian, Nasopharyngeal carcinoma, Lung, Bladder

## EXAMPLE 7

RNA Primer-initiated abortive transcription with *E. coli* RNA polymerase holoenzyme.

[0407] RNA Primer-initiated abortive transcription with *E. coli* RNA polymerase holoenzyme. *E. coli* RNA polymerase holoenzyme can initiate transcription from single-stranded DNA molecules lacking a promoter sequence. FIG. 31 shows an experiment in which denatured poly[dG-dC] (10  $\mu$ g/25 $\mu$ l reaction) was transcribed with *E. coli* RNA polymerase holoenzyme (1.9 pmoles/reaction). Abortive transcription was initiated with the dinucleotide GpC. GTP was the sole nucleoside-triphosphate available to elongate the primer. The other nucleoside-triphosphate encoded by the template strand (CTP) was omitted. FIG. 31a shows the presence of the trinucleotide product GpCpG and its dependence on GTP concentration. The results of FIG. 31b show that all of the detectable product is of one size, suggesting that omission of CTP effectively terminated transcription after the formation of the trinucleotide product.

[0408] *E. coli* RNA polymerase holoenzyme had a strong preference for bubble complex substrates over template strands that lacked a partially complementary non-template partner. FIG. 32 shows the relative transcriptional activities by *E. coli* RNA polymerase holoenzyme with a DNA bubble complex versus the corresponding single template strand. The RNA polymerase exhibited 70-fold higher levels of activity with Bubble complex 1 (FIG. 30b) than when it was provided with an equivalent molar amount of the template strand alone (FIG. 30c, FIG. 32, samples 1-4 versus samples 5-8). Similar results were obtained in experiments examining the preference of T7 RNA polymerase for bubble complex DNA.

[0409] RNA polymerases with diverse promoter recognition properties can use bubble complex 1 as a substrate for abortive transcription. FIG. 33 shows the results of an experiment in which Bubble complex 1 was incubated with *E. coli* holoenzyme, *E. coli* core RNA polymerase, phage T7 and phage SP6

RNA polymerases. The reaction buffer for *E. coli* holoenzyme and *E. coli* core polymerases included 150 mM Na-acetate. Na-acetate was omitted from the T7 and SP6 reactions because high salt concentrations inhibit these enzymes. All reactions contained 20 mM HEPES pH 8 buffer, 10 mM MgCl<sub>2</sub> and 2 mM DTT. The initiator ApA and UTP were each provided at 1 mM in all of the reactions. *E. coli* holoenzyme produced about 2-fold more product than *E. coli* core polymerase and about 10-fold more product per polymerase than the T7 and SP6 polymerases. The performance of the T7 and SP6 enzymes potentially could be improved with optimization of their reaction buffers.

[0410] The similar levels of activity by *E. coli* RNA polymerase holoenzyme and core enzyme suggests that the structure of the bubble complex is more important for recognition by the polymerases than the possession of a particular promoter sequence because the core enzyme lacks the promoter-sequence specificity that is characteristic of the holoenzyme. Similarly T7 and SP6 RNA polymerases lack the ability to recognize *E. coli* promoters.

[0411] Sensitivity of assays based on primer-initiated abortive transcription with radioactive precursors and autoradiographic detection. The sensitivity of a detection assay based on a primer-initiated abortive transcription reaction was determined by defining the minimal amount of Bubble complex 1 that could produce a detectable signal (FIG. 34). A series of abortive transcription reactions was performed with decreasing amounts of Bubble complex 1 (10 femtomoles -1 zeptomole/25  $\mu$ l reaction). Transcription was initiated with ApA and radioactive UTP. UTP was the only nucleoside triphosphate included in the reactions in order to limit the product to the trinucleotide ApApU. FIG. 34a and 34b show thin layer chromatographs of a time-course for each transcription reaction. Each TLC plate was exposed to X-ray film of 1 hour at -40C. After a 3 hour RNA polymerase abortive transcription reaction, signal from 10 femtomoles of bubble complex was clearly detectable and a faint signal from 1 femt mole of bubble complex was discernable in the original autoradiograph. An ApApU signal from 100 attomoles of Bubble complex 1

was detectable after 24 hours of transcription (FIG. 34b). The level of sensitivity can be improved with alternative methods of radioactive detection employing phosphorimaging or direct counting of beta emissions.

#### EXAMPLE 8

##### Incorporation of nucleotide analogs

[0412] The inherent ability of *E. coli* RNA polymerase holoenzyme to incorporate ATP analogs was tested initially with the T7A1 promoter which encodes a transcript with the 5' terminal sequence AUCGA. A list of the analogs is presented in Table 3. Radioactive tetramers were synthesized with the extension of the ApU initiator with CTP and radioactive GTP. Extension of the pentamer to a hexamer or longer would show that the ATP analog could be incorporated into the +5 position. Of 11 ATP analogs tested three were freely incorporated into the +5 position. 8-APAS-ATP produced almost exclusively the pentamer product acting as a strong chain terminator (Fig. 35, lane 2). Coumerin-ATP and  $\alpha$ -thio-ATP were freely incorporated producing full length radioactive transcripts (Fig. 35, lanes 9 and 12). Low yields of full length transcripts were produced in the presence of Tetramethylrhodamine-ATP and Lissamine-ATP (Fig. 35, lanes 5 and 7). Four analogs were not incorporated to a significant extent (Alexafluor-647-ATP, BODIPY-FL-ATP, Fluorescene-ATP and 3'-OMe-SF-ATP, Fig. 35, lanes 3, 4, 8 and 10). Two analogs appeared to inhibit transcription (Texas-Red-ATP and bis-3'-OMe-ATP, Fig. 35, lanes 6 and 11). bis-3'-OMe-ATP is composed of two 3'-OMe-ATPs linked together through a disulfide bond involving their 8-S groups. The reduced version of bis-3'-OMe-ATP, (3'-OMe-SH-ATP) was readily incorporated by *E. coli* RNA polymerase.

[0413] This initial round of screening employed 7 commercially available ATP analogs with the fluorescent groups attached to the 7 position of ATP via linker arms between 5-12 carbons long (Fig. 35, lanes 3-9).

[0414] Preliminary experiments with an APC based on the highly abortive T5 promoter DNA DG142 indicated that both coumarin-GTP and dinitrophenyl-GTP were efficiently incorporated into the +3 position. In these experiments radioactive ATP is incorporated into +4, after the GTP analog is incorporated into +3. In the reaction containing coumarin-GTP a radioactive, fluorescent tetramer was formed. In the experiment with DNP-GTP, a radioactive tetramer was produced with a retarded electrophoretic mobility.

- 169 -

Common Name	Vendor/ Maker	Catalog No.	Chemical / IUPAC name
Alexa Fluor 647 ATP	Molecular probes	A-22362	Adenosine 5'-triphosphate, Alexa Fluor 647 2'- (or 3')-O-(N-(2-aminoethyl)urethane), hexa(triethylammonium salt)
8-APAS-ATP	Dr. Michelle Hanna	-----	8-[4-azidophenacyl]thioadenosine 5'-triphosphate
BODIPY-FL-ATP	Molecular probes	A-12410	adenosine 5'-triphosphate, BODIPY® FL 2'-(or-3')-O-(N-(2-aminoethyl)urethane), trisodium salt (BODIPY® FL ATP)
Tetramethylrhodamine-6-ATP	PerkinElmer lifesciences, Inc.	NEL-440	N/A, ATP is linked to the fluorescent group through a linker of 6
Texas Red-5-ATP	PerkinElmer lifesciences, Inc.	NEL-441	N/A, ATP is linked to the fluorescent group through a linker of 5
Lissamine-5-ATP	PerkinElmer lifesciences, Inc.	NEL-442	N/A, ATP is linked to the fluorescent group through a linker of 5
Fluorescein-12-ATP	PerkinElmer lifesciences, Inc.	NEL-439	N/A, ATP is linked to the fluorescent group through a linker of 12
Coumarin-5-ATP	PerkinElmer lifesciences, Inc.	NEL-438	N/A, ATP is linked to the fluorescent group through a linker of 5
Coumarin-5-GTP	PerkinElmer lifesciences, Inc.	NEL-497	N/A, ATP is linked to the fluorescent group through a linker of 5
3'-OMe-SF-ATP	Dr. Michelle Hanna	-----	8-thioacetamidofluorescein-3'-O-methyl adenosine 5'-triphosphate
3'-OMe-SH-ATP	Dr. Michelle Hanna	-----	8-mercaptop-3'-O-methyl adenosine 5' triphosphate
Bis-3'-OMe ATP	Dr. Michelle Hanna	-----	8,8'-Dithiobis(-3'-O-methyl adenosine 5'-triphosphate)
$\alpha$ -thio-ATP	Trilink BioTechnologies	N-8005-1	Adenosine-5'-O-(1-thiatriphosphate)

Table 3.

## EXAMPLE 9.

## Primer Extension of P16 DNA with Various Enzymes, to incorporate SH-dUTP

[0415] A series of primer extension reactions were done in order to determine whether SH-dUTP could be incorporated into a DNA fragment using a variety of available enzymes. Before setting-up the reactions, di(8-UTP)disulfide was reduced to SH-dUTP by adding a large amount of DTT. The reduction was verified by comparing the absorbance of the treated bis-dUTP with an untreated sample (at 270nm and 330nm), using a spectrophotometer. Once this was completed, previously amplified P16 DNA was denatured with NaOH and annealed to P16DF2 (the reverse primer) by heating to 70°C then slow-cooling to room temperature. The DNA was then extended using Sequenase 2.0, Therminator Polymerase, Vent exo-, Klenow exo-, AMV Reverse Transcriptase and M-MLV Reverse Transcriptase. Each reaction was incubated for thirty minutes at the optimum temperature for the enzyme used. The final concentration of each dNTP was 200 µM, except for the dTTP+SH-dUTP reactions that consisted of 20 µM dTTP and 200 µM SH-dUTP. The samples were analyzed via 8% PAGE and after a 20 hour exposure the image seen in Figure 1 was produced. The results indicate Therminator Polymerase to be the most proficient at incorporating SH-dUTP in full-length products. It appears that the reactions done with 100% analog went to completion (Figure 36, lane 7), whereas the dTTP reactions appear to have ended prematurely at times (Figure 36, lanes 3,6,9,12,15 &18).

## EXAMPLE 10

## Primer Extension of P16 DNA with Therminator Polymerase to incorporate SH-dUTP

[0416] A series of primer extension reactions were done in order to verify that SH-dUTP could be incorporated into a full-length DNA fragment using Therminator Polymerase. Before setting-up the reactions, bis-dUTP was reduced to SH-dUTP by adding DTT to final concentration of 0.16 mM. The reduction was verified by comparing the absorbance of the treated bis-dUTP with an untreated sample (at 270nm and 330nm), using a spectrophotometer. Once this was completed, previously amplified P16 DNA was denatured with NaOH and annealed to P16DF2 (the reverse primer) by heating to 70°C then slow-cooling to room temperature. The DNA was extended in 100 µL reactions with final concentrations of 1X ThermoPol Buffer, 10 mM DTT, 2.75 pmoles P16 DNA, 11 pmoles P16DF2, 200 µM dATP, dGTP and dCTP (+<sup>32</sup>PdCTP), 200 µM SH-dUTP or 200 µM of dTTP or 20 µM ddTTP with 180 µM dTTP, with 4 units of Therminator Polymerase. Each reaction was incubated for four cycles of: 72°C for 30 minutes, 94°C for one minute and 59°C one minute. For each reaction, a 5 µL and a 20 µL sample were analyzed via 8% PAGE. After a 16 hour exposure the image seen in Figure 2 was produced. In the 20 µL SH-dUTP sample a second band is seen that was not visible in the first primer extension experiment (Figure 37, Lane 1). This could be due to the small volume that was analyzed, which would explain why the 5 µL SH-dUTP sample only exhibits one band (Figure 37, Lane 2). The second, slightly smaller band may be the result of unsuccessful SH-dUTP incorporation when the sequence requires three uridines in a row (first occurring at 169 bp). Unfortunately, the ddTTP reaction still did not form the thymine-terminating ladder needed for accurate size comparison (Figure 37, Lanes 5 & 6).

- 172 -

## EXAMPLE 11

### P16 Primer Extension Incorporating SH-dUTP with Therminator Polymerase

[0417] Further primer extension reactions were conducted using P16DF PCR as a template, oligonucleotide P16DF2 as the primer and Therminator Polymerase. The DNA was extended in 100  $\mu$ L reactions with final concentrations of 1X ThermoPol Buffer, 10 mM DTT, 2.75 pmoles P16 DNA, 11 pmoles P16DF2, 200  $\mu$ M dATP, dGTP and dCTP ( $^{32}$ PdCTP), 200  $\mu$ M SH-dUTP or 200  $\mu$ M of dTTP or 20  $\mu$ M ddTTP with 180  $\mu$ M dTTP, with 4 units of Therminator Polymerase in each reaction. Incubation was again done in four cycles of: 72°C for 30 minutes, 94°C for one minute and 59°C one minute. This time, the DTT used to reduced bis-dUTP was removed by centrifugal filtration (microconing) the samples after the extension reaction.

[0418] The samples were then treated with 5'IAF (5' iodoacetyl fluorescein), in the presence of 0.5 mM TCEP (tetracarboxylethyl phosphene). The treated samples analyzed on an 8% PAGE gel run at 1700V for approximately 2.5 hours. An AUTRAD was set-up (overnight), which indicated successful formation of full-length DNA in each of the reactions except the negative control (Figure 38, Lanes 1-4). Because only SH-dUTP was present in the first reaction, this supports the assumption that SH-dUTP is incorporated successfully by this method.

## EXAMPLE 12

[0419] Further experiments were performed with analogs. Clone DG142 is a degenerate variation of the phage T5N25 promoter which is recognized by the *E.coli* RNA polymerase and the DG142APC is the highly abortive bubble complex containing the T5N25 promoter region. As seen in Figure 40, the expected product: Lanes: 8. AUGA, 9 AUGcoumA, 10. AUcoumGA, 11. AUDnpGA, 12. AUG There is partial incorporation Coumarin ATP (lane 9). The incorporation of Coumarin GTP and DNP-GTP is shown in lanes 10 and

11. The radiolabel  $\alpha^{32}\text{P}$ -ATP can go in only after the incorporation GTP analog. Upon irradiation of the gel with UV, coumarin emission can be seen (Fig 40B). The fluorescent Coumarin GTP band matches with the radioisotope spot. The reactions 1-7 were carried out with TSP2361 complex and the transcript would be ACGCA. This complex was observed to be highly abortive, and no tetramer products were observed. Only trimer products were formed from this complex with the biotinylated initiator.

[0420] The following experiment was designed to make FRET pair PyMPO SpApU- Coumarin 5 GTP/ATP using the PCR amplified T7A1 promoter region from pAR1707. SpApU was made from SpA. From the TLC results it looked like 100% conversion of SpA to SpApU and PyMPO-SpApU was made from SpApU. PyMPO-SpApU was used as the initiator and Coumarin 5 GTP/ATP was incorporated to make a transcript of FRET pair. As seen in FIG 41. Evaluation of the FRET pair made by abscription using *E.coli* holo RNA polymerase enzyme and the promoter T7A1 as the template. The reaction was initiated by PyMPO-SpApU (lanes 4-8). For comparison ApU was also used as initiator (lanes 1-3). Coumarin ATP and GTP were the analogs used (lanes 3, 6, 7 and 8) and 3'OCH<sub>3</sub>ATP was as the terminator (lanes 3 and 8).

[0421] Comparing lane 1 and 4 it can be shown that PyMPO-SpApU can be used as an initiator and it can abscribe the 20mer product. Lane 2 shows the formation of tetramer AUCG (this lane has no ATP, so the expected product is only tetramer) and in lane 5 the tetramer is PyMPO-AUCG. The additional spot in that lane could not be explained. The above autoradiogram is the result of 36 hours of exposure. When the film was exposed for only 6 hours the additional spot appeared very light. In lane 3 and 8 the terminator 3'OCH<sub>3</sub>ATP was included. In lane 3 there is spot corresponding to the trimer AUC because the C is the radio labeled NTP. No product is observed, suggesting that the reaction is prevented from proceeding. For example, 3'OCH<sub>3</sub>ATP could not be competing for initiation. It is worth investigating

the status of 3'OC<sub>3</sub>ATP by TLC. Lane 7 shows a very faint spot corresponding to the product PyAUCGcouA.

[0422] The most interesting lane is 6 which show the FRET pair product (PyAUCCouGA). The radioactive NTP used in this lane  $\alpha^{32}$ P-ATP. So this can be incorporated only after the incorporation Coumarin G. There is an additional spot corresponding to PyAUCCouGAcouG. After that the abscription is abortive.

#### Incorporation mercapto CTP (HS-CTP) and then labeling with 5-IAF

[0423] This experiment was done to see whether HS-CTP can be incorporated during abscription and whether it can be labeled with 5-iodoacetamido fluorescein (5-IAF) after abscription. AU-HIP/S10 complex was used for this study. UpA was used as the initiator and incorporating mercapto CTP followed by  $\alpha^{32}$ P-UTP. The expected product is UAbisCU.  $\alpha^{32}$ P UTP can be incorporated only when the mercapto-CTP is incorporated first. The spot in lane 1 indicates the incorporation of mercapto-CTP followed by  $\alpha^{32}$ P UTP. The fluorescent spot in lane 1, 2 (after IAF) indicates that the incorporated mercapto CTP can be labeled with Fluorescein.

#### Incorporation of mercapto CTP (HS-CTP)

[0424] PCR product containing theT7A1 transcript was used as the template for this reaction. The incorporation of HS-CTP is shown by the band above the 20 mer band in panel A & B of FIG 11. Panel shows the full incorporation product (20mer+UAGC (HS-CTP)). Figure 11. Incorporation of HS-CTP during elongation by E.coli RNA polymerase holoenzyme. Autoradiogram of 8% polyacrylamide/7M urea is shown in (A) and that of 25% is shown in (B). Autoradiogram of TLC plate is shown in (C). RNA 20nt long complex (20 mer) was made from the T7 A1 promoter with initiator ApU (lane 1) and biotinylated ApU (lane 2 &3) and CTP, ATP and  $\square^{32}\bar{P}$  GTP in the presence

of heparin. Then it was elongated in the presence of UTP (lane 1, 2 & 3) and HS-CTP (lane 3). Samples for the TLC were treated with CIP for 1 hour at 37°C.

#### EXAMPLE 11

Evaluation of the analog 3'OMe-8-(2, 4-DNP-S)-ATP as a transcription substrate for *E. coli* RNA polymerase holoenzyme.

[0425] Incorporation of the analog 3'OMe-8-(2, 4-DNP-S)-ATP into RNA was tested by in vitro transcription with a DNA fragment containing the *E. coli* RNA polymerase-specific T7A1 promoter. Omission of UTP from an otherwise complete reaction mixture produces the 20-mer product:

[0426] AUC GAG AGG GAA CAC GGC GAA.

[0427] Inclusion of chain terminating ATP analogs causes production of the pentamer product AUCGA. All transcription reactions included the initiator ApU, CTP and radioactive GTP to allow the production of a radioactive tetramer AUCG (Figure 44, Lanes 1-5). Inclusion of ATP allowed extension of the tetramer to form the predicted radioactive 20-mer (Figure 44, Lane 2). Inclusion of either chain terminator 3'-OMe-ATP or 3'OMe-8-(2, 4-DNP-S)-ATP caused the production of a radioactive pentamer containing the corresponding analog (Figure 44, Lanes 4 and 5). Pentamer containing 3'-OMe-A had a faster electrophoretic mobility than the pentamer terminated by 3'OMe-8-(2, 4-DNP-S)-A (Figure 44, Lane 4 verses Lane 5).